

S t Nam Qu ry
side by side

Hit Count S t Nam
result set

DB=USPT,PGPB,JPAB,EPAB,DWPI; PLUR=YES; OP=OR

L6 fibrinogen adj2 binding and coagulase adj2 negative

3 L6

DB=USPT,PGPB,JPAB,EPAB,DWPI,TDBD; PLUR=YES; OP=OR

L5 fibrinogen adj2 binding and epidermidis

7 L5

DB=USPT,PGPB,JPAB,EPAB,DWPI; PLUR=YES; OP=OR

L4 (fbe or fig) adj5 gene\$ and staphylococc\$

0 L4

L3 (fbe or fig) adj5 gene\$ ans staphylococc\$

42895 L3

L2 (fbe or fig) adj5 gene\$

10 L2

L1 (fbe or fig) and epidermidis

0 L1

END OF SEARCH HISTORY

WEST[Help](#)[Logout](#)[Interrupt](#)[Main Menu](#)[Search Form](#)[Posting Counts](#)[Show S Numbers](#)[Edit S Numbers](#)[Preferences](#)[Cases](#)**Search Results -**

Terms	Documents
fibrinogen adj2 binding and coagulase adj2 negative	3

Database:

US Patents Full-Text Database
US Pre-Grant Publication Full-Text Database
JPO Abstracts Database
EPO Abstracts Database
Derwent World Patents Index
IBM Technical Disclosure Bulletins

S arch:

L6

[Refine Search](#)[Recall Text](#)[Clear](#)**Search History****DATE:** Tuesday, July 30, 2002 [Printable Copy](#) [Create Case](#)

WEST

Generate Collection

L3: Entry 20 of 25

File: EPAB

Oct 10, 1991

PUB-NO: DE003583987A1

DOCUMENT-IDENTIFIER: DE 3583987 A1

TITLE: TITLE DATA NOT AVAILABLE

PUBN-DATE: October 10, 1991

APPL-NO: DE03583987

APPL-DATE: October 31, 1985

PRIORITY-DATA: DE03583987A (October 31, 1985)

INT-CL (IPC): A61K 37/02; C07K 7/10; C12P 21/02

ABSTRACT:

A new antibiotic, designated epidermin (I), has the prim. structure: *Staphylococcus epidermidis* DSM 3095, which is resistant to (I), is new. (I) is made by aerobic cultivation of DSM 3095 at 34-37 deg.C on a complex nutrient soln. contg. 2-4% N source (e.g meat extract), 1-3% sugar or sugar alcohol; 0.25-1% alkaline earth carbonate and/or 0.25-0.5% alkaline earth hydroxide. The cells and inorganic salts are removed, then (I) isolated by (a) extracting with butanol at pH 8, evaporating the extract, dissolving the residue in MeOH and pptn. of lipids with ether or (b) adsorbing onto acrylic ester or polystyrene polymers, eluting with 99:1 MeOH-concn. H₂SO₄, neutralising with NH₃ and evaporating in vacuo. The isolate is then chromatographed on 'Sephadex LH-20'(RTM) to remove low mol.wt. peptides amino acid and salts, and subjected to liq-liq partitioning first in 3:1:3 n-butanol/ethyl acetate/0.1N acetic acid ((I) remaining at the starting position) and then in the neutral system 1:1 2-butanol/0.05N NH₄ acetate. Purified (I) is recovered as a colourless powder by freeze-drying.

elution with methanol/diluted hydrochloric acid, (c) the eluate is adjusted to a pH of 5.3 to 5.8, (d) the eluate is placed on a weak cation exchanger, (e) non-bound substances are subsequently washed out with a buffer solution at pH 7, (f) the active component is eluted out of the cation exchanger with a solution consisting of buffer substance, sodium chloride and methanol at pH 6.0 to 8.0 and for purification washed with water in order to remove salts and the epidermin is released from the resin with a methanol/acetic acid mixture and the solution is evaporated or freeze-dried, whilst the epidermin thus obtained may subsequently also be subjected to high performance liquid chromatography for extra purification.

CHOSEN-DRAWING: Dwg.0/10 Dwg.0/10

TITLE-TERMS: ISOLATE STAPHYLOCOCCUS CULTURE ADSORB POLYSTYRENE BASED
COPOLYMER ELUTION CHROMATOGRAPHY CATION EXCHANGE USEFUL ANTIBIOTIC TREAT
SKIN INFECTION

DERWENT-CLASS: A96 B04 D16

CPI-CODES: A04-B10; A04-C04; A12-M03; A12-V; A12-W11L; B02-E; B11-B; B11-C08D2; B12-A07;
D05-C02;

CHEMICAL-CODES:

Chemical Indexing M1 *01*

Fragmentation Code

H1 H100 H101 H181 H182 H4 H401 H481 H8 J0
J011 J012 J1 J171 J172 K0 L2 L250 M280 M311
M312 M313 M314 M315 M321 M331 M332 M333 M340 M342
M343 M349 M381 M391 M421 M510 M520 M530 M540 M620
M720 M903 N131 N161 Q233 V050 V901 V913 V923

Ring Index

63917

Registry Numbers

1327U 0502U

POLYMER-MULTIPUNCH-CODES-AND-KEY-SERIALS:

Key Serials: 0231 0306 3162 0418 1123 2020 2569 3264 3272 2769

Multipunch Codes: 014 034 04- 055 056 074 075 077 128 231 27& 473 53& 532 533
54& 623 624 642 645 720

SECONDARY-ACC-NO:

CPI Secondary Accession Numbers: C1990-006930

WEST



Generate Collection

L3: Entry 24 of 25

File: DWPI

Jun 10, 1997

DERWENT-ACC-NO: 1990-016158

DERWENT-WEEK: 199944

COPYRIGHT 2002 DERWENT INFORMATION LTD

TITLE: Isolating epidermin from staphylococcus epidermidis culture - by adsorption on styrene! based copolymer, elution and chromatography on cation exchanger, useful as antibiotic for treating skin infections

INVENTOR: FIEDLER, H; HOERNER, T; JUNG, G; KELLNER, R; WERNER, R; ZAEHNER, H;
FIEDLER, H P; HORNER, T; KELLNER, J R; WERNER, R G; ZAHNER, H

PATENT-ASSIGNEE:

ASSIGNEE

THOMAE GMBH KARL

CODE

THOM

PRIORITY-DATA: 1988US-0219698 (July 15, 1988)

PATENT-FAMILY:

PUB-NO	PUB-DATE	LANGUAGE	PAGES	MAIN-IPC
KR 9709289 B1	June 10, 1997		000	C07K001/14
EP 350810 A	January 17, 1990	G	018	
AU 8938103 A	January 18, 1990		000	
PT 91171 A	February 8, 1990		000	
DK 8903506 A	January 16, 1990		000	
JP 02084194 A	March 26, 1990		000	
ZA 8905362 A	March 27, 1991		000	
EP 350810 B1	September 29, 1993	G	021	C07K001/14
DE 58905744 G	November 4, 1993		000	C07K001/14
ES 2059645 T3	November 16, 1994		000	C07K001/14
IE 62402 B	January 25, 1995		000	C07K001/14
CA 1336896 C	September 5, 1995		000	C12P021/02
JP 2777205 B2	July 16, 1998		012	C12P021/02

DESIGNATED-STATES: AT BE CH DE ES FR GB GR IT LI LU NL SE AT BE CH DE ES FR GB GR IT LI
LU NL SE

CITED-DOCUMENTS:1.Jnl.Ref; A3...9139 ; EP 181578 ; EP 27710 ; No-SR.Pub

APPLICATION-DATA:

PUB-NO	APPL-DATE	APPL-NO	DESCRIPTOR
KR 9709289B1	July 15, 1989	1989KR-0010095	
EP 350810A	July 7, 1989	1989EP-0112446	
JP02084194A	July 14, 1989	1989JP-0182372	
ZA 8905362A	July 14, 1989	1989ZA-0005362	
EP 350810B1	July 7, 1989	1989EP-0112446	
DE58905744G	July 7, 1989	1989DE-0505744	
DE58905744G	July 7, 1989	1989EP-0112446	
DE58905744G		EP 350810	Based on
ES 2059645T3	July 7, 1989	1989EP-0112446	
ES 2059645T3		EP 350810	Based on
IE 62402B	July 14, 1989	1989IE-0002283	
CA 1336896C	July 14, 1989	1989CA-0605673	
JP 2777205B2	July 14, 1989	1989JP-0182372	
JP 2777205B2		JP 2084194	Previous Publ.

INT-CL (IPC): A61K 0/00; C07G 11/00; C07K 1/14; C07K 3/12; C07K 7/10; C07K 15/04; C07K 17/10; C12N 11/08; C12P 1/04; C12P 21/02; C12R 1/45; C12P 21/02; C12R 1/45; C12P 21/02; C12R 1/45

ABSTRACTED-PUB-NO: EP 350810A

BASIC-ABSTRACT:

the polypeptide antibiotic epidermin (I) is isolated and purified from a culture of a *Staphylococcus epidermidis* strain by (a) applying the culture broth or filtrate to a styrene-divinyl copolymer (A), (2) eluting active ingredients with MeOH-dil HCl, (3) adjusting eluate to pH 5.3-5.8; (4) applying to a weak cation exchanger (B) (5) washing-out non-bound cpds. with pH7 buffer, (6) eluting (I) with pH 6-8 buffer contg. and MeOH, (7) readsorbing (I) onto (A) washing the resin with water (desalting) and eluting with MeOH-MeCOOH mitd., (8) evaporating or freeze-drying the eluate, and opt. (9) further purifying by h.p.l.c.

The *S. epidermidis* strains used are pref. DSM 3095 or NC18 11536.

USE/ADVANTAGE - (I) is known for treatment of skin infections such as eczema, impetigo, cellulitis and acne. This method is simple and produces significantly higher yields of (I) then known processes.

ABSTRACTED-PUB-NO:

EP 350810B

EQUIVALENT-ABSTRACTS:

Process for isolating epidermin from a culture broth or a culture filtrate of a strain of *Staphylococcus epidermidis* and for purifying this substrate, characterised in that (a) the culture filtrate or culture broth is added to a styrene-divinyl copolymer, (b) the active component is released from the resin by

A:Reference number: S41539; MUID:94224142
 A:Accession: S41539
 A>Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1933 <MCD>
 A:Cross-references: EMBL:Z18852; NID:9397525; PIDN:CAA79304.1; PID:9397526

Query Match 14.6%; Score 452; DB 2; Length 933;
 Best Local Similarity 26.6%; Pred. No. 3; 1e-12;
 Matches 154; Conservative 105; Mismatches 234; Indels 86; Gaps 26;

```

QY 8 SDEKNDVNNQSIINTDNNQIIRKKEETNNYDGIEKRESDTESTNVNDEATFLAK 67
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 51 SNEKSDSSVSAAPRTDITN-----SDTKTSNTN---NGETSVAO 91
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 68 TPQDNTLTHEEVKSSSVSSSSIDTAQPSHTTINREESVOTSDVSHVSPANS 127
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 92 NP-----AQDETQSSSTNATTEETPTVTEATTTTNOANTPATQOS-SNTAAELVN- 143
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 128 KIKESNTEGKEENTIEQPNKVEDSTTSOPSGYTNIDEKISNODELLN-LPINEYENNA- 136
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 144 --QTSNETTFNDITV-----SSVNSPQNSTNAENSTTODTTEATPSN--NES 189
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 187 RPLSTTSAPSIKRYTVN-----QLAAEQGSNVNHLIKYTDQ--SITEGYDSEGV 235
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 190 APQSTDAENKRVNQAANTSAFPMRAFSLAAVADAPAGDITNQLNTVYGI-DSGTT 248
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 236 IKANDAENLIYDVFEVDKYSQDITMVIDIKNTVPSDLTSTTIKIDNNGEIIATG 295
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 249 VYPIQAGYVKNLYGFSVPSNSAVKGDTEKITVPKELNLNGYSTAKVPPIMAGD-QVLANG 307
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 296 TYDKNKNOITFTFDYDVKENIKAKHLKLTYSIDKSKVPNNNTKL-DVEYKTALS--VN 352
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 308 VIDS-DCNVYITFTDYNTKDVAKATLMPAYID-----PEVKKTKGVTLATIGSTTAN 362
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 353 KTIYEVORPENNTANLQSMETNIDKKNHTVEQTIYINPL--RYSAKETNVNISGNGDE 410
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 363 KTVLDYEEKYKRFYKLSIKGTIDQIDKTNNTYRGTIYVNPFGDVNIAFVLTGKPKPTDS 422
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 411 GSTIIDSTIIKVKYGVGNQNLPPDSNRIYDSEVEDYTNNDYADLGNNDVINIFG---- 466
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 423 NALIDQQTISIKYKVDNAADLSEYFV-NPENEDEDVNSYNTFFPNNOYKVEFNTPPD 481
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 467 NIDSPYIKYISKYDPN-KDDYTTIQOTVMTQTTINEYTCG--FRFASDYNTIAFSTSSG 523
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 482 OTTPIYIYVNGHIDPNKGD-----LALRSTLYGYNSNIIMRSMWDEVAFFNNGSG 534
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 524 QGQG-DLP--PEKTYKIGDY--VWEDVDKD-GIONTNDN 556
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 535 SGGDIDKPVVPEQDPPEIEPIPEDSDSDPSGSDS 573
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

```

Query Match 15.9%; Score 490; DB 2; Length 953;
 Best Local Similarity 28.2%; Pred. No. 7; 7e-14;
 Matches 162; Conservative 83; Mismatches 236; Indels 94; Gaps 22;

```

QY 42 GIEKRESDTESTNVNDEATFLAKTPQDNTLTHEEVKSSSVSSSSIDTAQPSH 101
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 45 GHEAKAAEHTNGELNQSKEET---APSEMK--TTEKV--DSROLKNDTQATADQPKV 96
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 102 TTINREESVOTSDVSHVSPANSKIKESNTEGKEENTIEQPNKVEDSTTSOPSGY 161
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 97 T-----MPSDATVKETSSNMOS-----PQNTASOSTTQTSNV 129
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 162 TNIDEK---ISNODELLNLPINEYEN-KARPLSTSAOPSIKRYTVNQLAA-EQGSNVN 216
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 130 TTNDKSSSTTYSNETKSNL--TQAKNVSTTPKTTTIKQALNKAQVNTVAAPQSGTINVND 187
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 217 LIKVTDOSI-----TEGYDSEGYIKAHDAENLIYDVFEVDKYSQDITM 263
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 188 KYHFINIDIALDKGVNKTGTGTEFWATSSDYLK-----LKANYTIDDSVEKEGDFET 239
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 264 VIDKNTVPSDLTDSFTTIKIDNNGEIIATGTYDNKKQITVFTTQVYDKYENIKAKHLK 323
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 240 FRYGGYFRGGSVRLPSQTONLNAQGNIIAKGIYDKTNTTYYTTNTVDDQTTNVSSE 299
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 324 LRSYIDKSVPPNNNTKLDEYKTAALSVKNTTVEYORPENNTANLQSMETNIDTKNHT 383
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 300 QVAFKAKRENATDKTAYKKEVTLGNDYTSKQVIVDY---GNKGGQQLLSSTYINNEELS 356
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 384 VEOITVIN-PLRYSAKETNV-NISGNGEGSTIIDSTIIKVKYGVGNQNLPPDSNRIYDY 441
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 357 RNMITYVNPQPKTYKTEFVNTLT-----GYKFPNDAKKFKIYEVTDOQNFVDS-FTPD 410
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 442 SEVEDYTN-DYAOLGNNNDVINIFGN---IDSPYIIKVISKYPNNDYTTIQOTYTM 496
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 411 SKLKDVTFQFDVYSNDNKTATVLDLNGOSSDKQYIIIOQVAVPNNSTDNKIDYTTLET 470
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 497 OTTINEYGEFTFASDYNTIAFSTSSGQGGDLPEPEKTKIGEVWEDVDKGIQNTNDN 556
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 471 QNGKSSWSN-----SYSNVNGSSTANGD-----QKKYNLGDYVWEDTNKKGKQDA--N 516
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 557 EKPLSNVLTLYTPDQTS-KSVRTDEDKRYQPDGV 590
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 517 EKIGKGVYIILKDSNGKELDRTTDBNGKYQFTGL 551
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

```

RESULT 7
 S41539
 fibrinogen-binding protein - Staphylococcus aureus
 N:Alternate names: clumping factor
 C:Species: Staphylococcus aureus
 C:date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 15-Oct-1999
 C:Accession: S41539; S36630
 R:McDevitt, D.; Francois, P.; Vaudeaux, P.; Foster, T.J.
 Mol. Microbiol. 11, 237-248, 1994
 Title: Molecular characterization of the clumping factor (fibrinogen receptor) of Staphylococcus aureus

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 30, 2002, 10:28:50 ; Search time 85.16 Seconds
(without alignments)
773.447 Million cell updates/sec

Title: US-09-147-405-11
Perfect score: 3087
Sequence: 1 HHHHHHPSDEKNDVINNN.....SKSVRTDEDKYQPDGVQVD 593

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues
Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: A_Geneseq_032802.*
2: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1980.DAT:*
3: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1981.DAT:*
4: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1982.DAT:*
5: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1983.DAT:*
6: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1984.DAT:*
7: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1985.DAT:*
8: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1986.DAT:*
9: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1987.DAT:*
10: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1988.DAT:*
11: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1989.DAT:*
12: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1990.DAT:*
13: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1991.DAT:*
14: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1992.DAT:*
15: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1993.DAT:*
16: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1994.DAT:*
17: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1995.DAT:*
18: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1996.DAT:*
19: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1997.DAT:*
20: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1998.DAT:*
21: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1999.DAT:*
22: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2000.DAT:*
23: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:*
24: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2002.DAT:*
25: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2003.DAT:*
26: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2004.DAT:*
27: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2005.DAT:*
28: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2006.DAT:*
29: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2007.DAT:*
30: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2008.DAT:*
31: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2009.DAT:*
32: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2010.DAT:*
33: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2011.DAT:*
34: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2012.DAT:*
35: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2013.DAT:*
36: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2014.DAT:*
37: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2015.DAT:*
38: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2016.DAT:*
39: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2017.DAT:*
40: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2018.DAT:*
41: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2019.DAT:*
42: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2020.DAT:*
43: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2021.DAT:*
44: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2022.DAT:*
45: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2023.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3015	97.7	1092	19	Staphylococcus epi
2	2812.5	91.1	991	21	Cell wall protein
3	2812.5	91.1	991	21	Staph. epidermidis
4	1146.5	37.1	1166	20	S. aureus SdrE pro
5	857	27.8	278	22	S. epidermidis ope
6	576	18.7	1315	20	S. aureus SdrD pro
7	576	18.7	1349	22	Staphylococcus aur
8	576	18.7	1349	22	Staphylococcus aur
9	506	16.4	1802	21	Cell wall protein
10	506	16.4	1802	21	Staph. epidermidis
11	504	16.3	1155	22	S. epidermidis ope

12	500.5	16.2	932	22	AAU34082	Staphylococcus aur
13	500.5	16.2	932	22	AAU36845	Staphylococcus aur
14	456	14.8	841	22	AAU34283	Staphylococcus aur
15	456	14.8	841	22	AAU37158	Staphylococcus aur
16	456	14.8	930	20	AAU08641	S. aureus SdrC pro
17	454	14.7	936	18	AAU89801	Staphylococcus aur
18	452	14.6	933	21	AAU58435	Staphylococcus aur
19	452	14.6	933	22	AAU69508	Staphylococcus aur
20	449	14.5	1021	22	AAU33975	Staphylococcus aur
21	449	14.5	1021	22	AAU36951	Staphylococcus aur
22	401.5	13.0	767	20	AAU08640	S. aureus ClfB pro
23	376	12.2	767	22	AAU34403	Staphylococcus aur
24	376	12.2	767	22	AAU37547	Staphylococcus aur
25	373	12.1	940	11	AAU07070	Fibrinogen-binding
26	359.5	11.6	238	18	AAU28019	Staphylococcus aur
27	349	11.3	345	19	AAU31555	Fibrinogen-binding
28	319.5	10.3	978	22	AAU33960	Staphylococcus aur
29	319.5	10.3	1001	22	AAU37093	Staphylococcus aur
30	309.5	10.0	1018	22	AAU34301	Staphylococcus aur
31	309.5	10.0	1018	22	AAU37245	Staphylococcus aur
32	304	9.8	1027	18	AAU89806	Staphylococcus aur
33	301	9.8	1018	9	AAU2115	Fibrinogen binding
34	259.5	8.4	1112	20	AAU08603	S. pyogenes SFBP-
35	219	7.1	537	7	AAU60452	Sequence of the As
36	214	6.9	1279	22	AAU83047	S. epidermidis ope
37	213.5	6.9	2206	21	AAU18254	Plasmodium falcipa
38	208.5	6.8	251	18	AAU89804	Staphylococcus aur
39	206.5	6.7	2368	22	AAU34139	Staphylococcus aur
40	206.5	6.7	2368	22	AAU36796	Staphylococcus aur
41	204.5	6.6	2727	22	AAU61674	Drosophila melanog
42	204	6.6	5024	22	AAU82935	S. epidermidis ope
43	202.5	6.6	1997	21	AAU18287	Plasmodium falcipa
44	202	6.5	1308	21	AAU18167	Plasmodium falcipa
45	201	6.5	865	21	AAU18278	Plasmodium falcipa

ALIGNMENTS

RESULT	1	
AAU41602	AAU41602 standard; Protein; 1092 AA.	
AC	AAU41602;	
XX	22-JUN-1998 (first entry)	
XX	Staphylococcus epidermidis fibrinogen binding protein FIG.	
XX	Fibrinogen binding protein; FIG: aggregation; infection;	
XX	coagulase-negative Staphylococcus; therapy; diagnosis;	
XX	Immunisation; Immunogen; vaccine.	
OS	Staphylococcus epidermidis strain HB.	
XX	Key	Location/Qualifiers
XX	Peptide	1..51
FT	Protein	/label= Sig-peptide
FT	Region	52..1092
FT	Region	/label= Mat-protein
FT	Region	52..824
FT	Region	/note= "non-repetitive region; harbours
FT	Region	fibrinogen binding activity"
FT	Region	825..1040
FT	Region	/note= "Asp-Ser dipeptide repeat region"
FT	Region	1053..1057
FT	Region	/note= "cell wall anchoring motif"
XX	MO9748727-A1.	
XX	24-DEC-1997.	
PD	18-JUN-1997;	97WO-SE01091.
XX		
PF		

QY 308 FTYVDKXENIKAHKLTISYIDSKYPPNNNTKIDVEYKTAALSSVNTKTIIVEYORENENRT 367
DB 405 ftyvdkyenikahkltisydskypnnntkldveyktalssvntkltiveyqkpenent 464
QY 368 ANQSMFTNIDTAKNHTVEQTIYINPLRYSAKETNVAISNGDGSSTIIDSTIKYKYK 427
DB 465 anqsmftndctknhtveqtiynplrysakentvnisngdgsstiidstlikykkyg 524
QY 428 DNGNLPSNRNIYDYSEYEDVTNDYALQGNNDVNIINFGNIDSPYIKYISKYDPKDDY 487
DB 525 dngnlpsnrniydyseyedvtndyaglggnndvniinfgnidspylklyskypdkddy 584
QY 488 TTIIDVVTMOTINEXTEGERRASNDNTAFSSSGOGGDDPREKTYIKGVWEDVVK 547
DB 585 ttiidvvtmotinextegertrasndntafstssgqgddlprektyikgvdvdk 644
QY 548 DGIQNTNDNEKPLSNVLTLYTPDGTSGSVRTDEDEKYOFGVO 591
DB 645 dgiqntndnekplsnvltlytpdgtsgsvrtdeegkyqfvgik 688

SULT 4
AY08643
ID AAY08643 standard; Protein; 1166 AA.
XX AAY08643;
AC
XX
DT 09-AUG-1999 (first entry)
XX
DE S. aureus Sdrc protein.
XX
XX Fibrinogen-binding protein; alpha chain; beta chain; Clfb; Sdrc; Sdrc;
KM Sdrc; fibrinogen; medical device; competitive inhibitor; pharmaceutical;
KM treatment; infection; septicemia; osteomyelitis; mastitis; endocarditis;
KM extracellular matrix; vascular graft; vascular stent; vaccine;
KM intravenous catheter; artificial heart valve; cardiac assist device;
XX antibacterial.
OS Staphylococcus aureus.
XX
XX WO927109-A2.
XX
XX 03-JUN-1999.
XX
XX 25-NOV-1998; 98WO-US25246.
XX
XX 31-AUG-1998; 98US-0098427.
PR 26-NOV-1997; 97US-0066815.
XX
XX (EIDH/) EIDHIN D N.
XX (FOEF-) FORPAS T A. BIORESEARCH IRELAND.
XX (FOST/) FOSTER T J.
XX (HOOK/) HOOK M A O.
XX (INH/) INHIBITEX INC.
XX (JOSE/) JOSEFSSON E.
XX (PAT/) PATI J M.
XX (PERK/) PERKINS S E.
XX
PI Eldhin DN, Foster TJ, Hook MAO, Josefsson E, Patti JM;
PI Perkins SE;
DR MPI: 1999-357844/30.
DR N-PSDB: AAX71594.
XX
XX Staphylococcus aureus fibrinogen-binding proteins for treating
PT septicemia, osteomyelitis, mastitis or endocarditis
XX
XX Claim 8; Fig 9; 143pp: English.
XX
XX This invention describes novel Staphylococcus aureus fibrinogen-binding
CC proteins that bind both the alpha and beta fibrinogen chains. The
CC proteins (and their encoding nucleic acids are Clfb, Sdrc, Sdrc and
CC Sdrc). Staphylococcus aureus is thought to utilize fibrinogen to adhere

CC to medical devices, binding proteins that bind both the alpha and beta
CC fibrinogen chains (Clfb, Sdrc, Sdrc and Sdrc) can therefore be used as
CC competitive inhibitors to block this binding. Antibodies against Clfb,
CC Sdrc, Sdrc and Sdrc inhibit Clfb, Sdrc, Sdrc and Sdrc mediated binding.
CC The proteins of the invention can be used in a pharmaceutical composition
CC for the treatment of Staphylococcus aureus infection e.g. septicemia,
CC osteomyelitis, mastitis or endocarditis or to inhibit the binding of
CC S. aureus to the extracellular matrix. The proteins or their fragments
CC may be used to coat a medical device to reduce the S. aureus infection of
CC an indwelling medical device, especially where the medical device is
CC selected from the group consisting of vascular grafts, vascular stents,
CC intravenous catheters, artificial heart valves, and cardiac assist
CC devices. Clfb, Sdrc, Sdrc, Sdrc, or an active fragment, subdomain or
CC encoding gene may be used as a vaccine. The DS (aspartate serine) repeat
CC region or a gene encoding it may be used as an identifying probe for the
CC identification of genes and encoding proteins from Staphylococcus aureus
CC (other than Clfb), S. hemolyticus, S. lugdunensis, and S. schleierferti.
CC The proteins of the invention have antibacterial activity.
XX
SQ Sequence 1166 AA:
QY
XX
XX Query Match 37.1%; Score 1146.5; DB 20; Length 1166;
Best Local Similarity 42.5%; Pred. No. 6.4e-54;
Matches 263; Conservative 110; Mismatches 191; Indels 55; Gaps 18;
QY 8 SSDEKNDVYINNQSINTDNNQI--IKKETNNNDGIEKRSERSTETNDENATPL 65
DB 66 atsdnkevseletennstennstnpkke--tndsqpeakkeststsqgnvntat 123
QY 66 OKTPDONTLHEEYKES---SSSESSNSIDAOQPSHTINREESVQTSQVDESHS 122
DB 124 tekpgn--tekenvpnsdktatetvyleekapnnt--nndvltkps-----ts 172
QY 123 DRANSKIKESNT---ESGREENTIEQPNVKEDSTTSQPSGYTNIDEKIS-NODELLNP 178
DB 173 epstselqkptpgeestnlensqbpptskvd---nqytdanepkpnvskelkmp 229
QY 179 -----INYEKARPLSTTSQAOPSTKR-----TYNQLAEQSNVHLIVTQD 223
DB 230 ekikelvndnsrldhsktpvatpavapkrvakmrfavaqpaavaasnvdllkvtkx 289
QY 224 SITEQDSEGYKADAEHLIDYVTFEVDKVKSGDITMTVDIDKNTVPSDLTFTIPK 283
DB 290 tikvg-dgkdnvaahdgkdieydeftldnkvkksqtdmtindkvnlpbdldkndpid 348
QY 284 IKDNGEIIATGYDNKNKQITTYTDDYDKYENIKAHKLTYSYIDSKYPPNNNTLDVE 343
DB 349 ltdpsgevlakgftkackqlytftcdydkyediksriltlysidkkvyp-netalnt 407
QY 344 YKTALSSVAKTIIVEYORENENTANLOSFTNIDTNRKNTVEQTIYINPLRYSAKETNVA 403
DB 408 fatagketeqnvltvdyqdmvngdsnlqslftldedkqtleqglyvnpkksaltukvd 467
QY 404 ISNG-----DEGSTIIDSTIKYKYVGDNDNLPSNRNIYDYSEYEDVTND-DYA 453
DB 468 lagsqvdgynklngstldqnteklykvnssdqglpqsnnlyfsgyedtsgfdnk 527
QY 454 QLGNNNDVNIINFGNIDSPYIKYISKYDPKDDYTTIIOQTVMOTINEXTEGERRASTD 513
DB 514 NTIAFTSSGOGGDD-PEKTYIKIDYVWEDVDRKGIQNTNDNEKPLSNVLTLYTPDG 572
QY 586 nftvnsndcgygqgvtkpeklykigdvwedvdkvgqtdexkexmanvltlytpdg 645
QY 573 TSKSVRTDEDEKYOFGVO 591
DB 646 tkksvrtdeanghyefgik 664
RESULT 5
AAG82803

ID AAG82803 standard; Protein: 278 AA.
 XX AAG82803;
 XX
 DT 03-SEP-2001 (first entry)
 XX
 DE S. epidermidis open reading frame protein sequence SEQ ID NO:2700.
 XX
 KM Staphylococcus epidermidis SRI strain; infection: diagnosis;
 KM vaccination; endocarditis.
 OS Staphylococcus epidermidis.
 XX
 PN WO200134809-A2.
 XX
 PD 17-MAY-2001.
 XX
 PF 09-NOV-2000; 2000WO-US10782.
 XX
 *X 09-NOV-1999; 99US-0164258.
 PA (GLAX) GLAXO GROUP LTD.
 XX
 FI Kimmberly MJ;
 XX
 DR WPI; 2001-316495/33.
 DR N-PSDB; AAH53653.
 XX
 PT Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
 PT useful for vaccinating against infections, e.g. endocarditis -
 XX
 PS Claim 18; Page 705; 2188pp; English.
 XX
 XX AAH53970 represent nucleic acids (I) encoding polypeptides
 CC (II), given in AAG81454 to AAG83120, from Staphylococcus epidermidis.
 CC (I) and (II) can have antibacterial activity and therefore can be used
 CC in vaccination. The nucleic acids (I) may be used to produce the
 CC S. epidermidis polypeptides (II) via the production of vectors
 CC containing them which are used to produce hosts cells which express the
 CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
 CC used to vaccinate subjects and to raise antibodies against the bacteria.
 CC The polypeptides may also be used to assay for other inhibitors of their
 CC activity and therefore identify compounds that may be used for the
 CC treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
 CC AAH55090 represent specifically claimed S. epidermidis genomic DNA
 CC polynucleotide sequences from the present invention. AAH55091 to
 CC AAH55098 represent oligonucleotide sequences and primers which are used
 CC in the amplification of the present invention.
 CC N.B. The present invention specifically claims all the polynucleotide
 CC sequences given in the sequence listing of the present specification,
 CC however the sequence listing only goes up to SEQ ID NO:4454 so even
 CC though sequences are given in the disclosure for SEQ ID NO:4465 to 4472,
 CC no sequences are present for SEQ ID NO:4455 to 4464.
 CC
 XX
 XX Sequence 278 AA:
 SQ
 Query Match 27.8%; Score 857; DB 22; Length 278;
 Best Local Similarity 93.0%; Pred. No. 4,4e-39;
 Matches 172; Conservative 3; Mismatches 8; Indels 2; Gaps 1;

QY 188 PLST 192
 II I
 Db 253 klsvt 257
 RESULT 6
 AAY08642
 ID AAY08642 standard; Protein: 1315 AA.
 XX
 AC AAY08642;
 XX
 DT 09-AUG-1999 (first entry)
 XX
 DE S. aureus SdrD protein.
 XX
 XX Fibrinogen-binding protein; alpha chain; beta chain; ClfB; SdrC; SdrD;
 KM SdrE; fibrinogen; medical device; competitive inhibitor; pharmaceutical;
 KM treatment; infection; septicemia; osteomyelitis; mastitis; endocarditis;
 KM extracellular matrix; vascular graft; vascular stent; vaccine;
 KM intravenous catheter; artificial heart valve; cardiac assist device;
 KM antibacterial.
 KM
 XX Staphylococcus aureus.
 XX
 OS
 XX WO9927109-A2.
 XX
 PD 03-JUN-1999.
 XX
 PF 25-NOV-1998; 98WO-US25246.
 XX
 PR 31-AUG-1998; 98US-0098427.
 PR 26-NOV-1997; 97US-0066815.
 XX
 PA (EIDH/) EIDHIN D N.
 PA (FOF/) FORFAS T/A BIORESEARCH IRELAND.
 PA (FOST/) FOSTER T J.
 PA (HOOK/) HOOK M A O.
 PA (INHIT-) INHIBITEX INC.
 PA (JOSE/) JOSEFSSON E.
 PA (PATV/) PATTI J M.
 PA (PERK/) PERKINS S E.
 XX
 PI Eidlun DN, Foster TV, Hook MAO, Josefsson E, Patti JM;
 PI Perkins SE;
 XX
 DR WPI; 1999-357844/30.
 DR N-PSDB; AAX77593.
 XX
 PT Staphylococcus aureus fibrinogen-binding proteins for treating
 PT septicemia, osteomyelitis, mastitis or endocarditis
 XX
 PS Claim 8; Fig 8; 143pp; English.
 XX
 CC This invention describes novel Staphylococcus aureus fibrinogen-binding
 CC proteins that bind both the alpha and beta fibrinogen chains. The
 CC proteins (and their encoding nucleic acids are ClfB, SdrC, SdrD and
 CC SdrE). Staphylococcus aureus is thought to utilize fibrinogen to adhere
 CC to medical devices, binding proteins that bind both the alpha and beta
 CC fibrinogen chains (ClfB, SdrC, SdrD and SdrE) can therefore be used as
 CC competitive inhibitors to block this binding. Antibodies against ClfB,
 CC SdrC, SdrD and SdrE inhibit ClfB, SdrC, SdrD and SdrE mediated binding.
 CC The proteins of the invention can be used in a pharmaceutical composition
 CC for the treatment of Staphylococcus aureus infection e.g. septicemia,
 CC osteomyelitis, mastitis or endocarditis or to inhibit the binding of
 CC S. aureus to the extracellular matrix. The proteins or their fragments
 CC may be used to coat a medical device to reduce the S. aureus infection of
 CC an indwelling medical device, especially where the medical device is
 CC selected from the group consisting of vascular grafts, vascular stents,
 CC intravenous catheters, artificial heart valves, and cardiac assist
 CC devices. ClfB, SdrC, SdrD, SdrE, or an active fragment, subdomain or
 CC encoding gene may be used as a vaccine. The DS (aspartate serine) repeat
 CC region or a gene encoding it may be used as an identifying probe for the
 CC identification of genes and encoding proteins from Staphylococcus aureus

CC (other than Cifa), *S. hemolyticus*, *S. lugdenensis*, and *S. schlerferi*.
 CC The proteins of the invention have antibacterial activity.

XX Sequence 1315 AA;

Query Match 18.7%; Score 576; DB 20; Length 1315;
 Best Local Similarity 30.1%; Pred. No. 5.3e-23;
 Matches 186; Conservative 112; Mismatches 228; Indels 92; Gaps 29;

QY 8 SSDEKNDV---INNOSINTDNNQIKKEETNNYDGIKRSRDEPSTTNVDE-NEAT 63
 Db stklneatsasdgskdvdmqnglndkndngkemsqngnettsngkllleke 114
 QY 64 FLOKTPQDNHILTEEEKSESSVSSNSITDAQPSHTTINRESVQTSNVEDSHVSD 123
 Db svsgltgknkvestaksdeqspkstnedlntkq-----tlsngealq-pdlgenksvvn 168
 QY 124 FANSKIKESNTSGKEENTIEQPNKVKEDSTTSQPSGYTNI-DEKISMODELL--NLPIN 180
 Db -----vgptneenkvd-----aktestt-----lnksdalksndeclvdmnsns 209
 QY 181 EYENKARPLSTSAQP---SIKRVYVNLQALAEQGSVNVNHLIKVTDOSTEGYDSESVI 236
 Db nneennadllpkstapkrlntrmlaavqpssteaknvnlltanltlvddaknmkiv 269
 QY 237 KAHDAENLIYDVFEVDVVKYSGDTMNVVDIDKNV-----PSDLTDSFTIPKIKD-NSG 289
 Db padgylslksqilt--vddkxksqdyftiky-sdlvqyglnpedlkn--lgdlkdpng 323
 QY 290 EIIATGYDNKNKOITTFDYVDKYEENIKAKHLKTSYIDKSKVPNNNTKLDVEYKTALS 349
 Db etatahdcanmllytfcdyvrfrsvqmglnyslymdcltp--vskndvefnvtlig 381
 QY 324 SVAKTIVEVQRPN--ENKRAVLOSMTNIDTKNHT-----VEQTIYINPLRYSAKE 399
 Db 382 nttktlanlqypdyvvnknsqsaft--etvshvgnkepfgyyktllynpensnltln 439
 QY 400 TNVNI-----SGNGDEGSTIIDSTIIIVYKVGQGNQNPDSRIYDYS--EDEDYTNDDYA 453
 Db 440 aklvqgahsyapnmlqinkdvrdlkiyvpkytl--nkgydvntkeltvtn-qyl 495
 QY 454 Q---LGNNDVNINFGINDSPYIIKIVISKYDPNNDYTTIQOTVMTQITINETYGEFRYA 510
 Db 496 qkltygdmsaavldfgnadaayvwmvntkfgynsepslvtqmatltsign-----ksv 549
 QY 511 SYDNTIAFSTSSGOGDLPPEKTYRTIGDYVMEVDKDGIONTNDNKPISNVLTLYTP 570
 Db 550 stgnalqftnngsgag-----qevykignyvwedtnknvgel--gekgygnvltv-fd 602
 QY 571 DGTSKSVR---TDDEGKY 585
 Db 603 nnttkygeavtkedgsy 620

RESULT 7
 AA034402 standard; Protein; 1349 AA.
 ID AA034402;
 XX

AC AA034402;
 DT 14-FEB-2002 (first entry)
 XX

DE Staphylococcus aureus cellular proliferation protein #678.
 XX

KW Antisense; prokaryotic cellular proliferation protein;
 XX antibiotic; antibacterial; drug design.
 OS Staphylococcus aureus.
 PN WO200170955-A2.
 XX
 PD 27-SEP-2001.

XX 21-MAR-2001; 2001WO-US09180.
 PF

XX 21-MAR-2000; 2000US-191078P.
 PR 23-MAY-2000; 2000US-206848P.
 PR 26-MAY-2000; 2000US-207727P.
 PR 23-OCT-2000; 2000US-242578P.
 PR 27-NOV-2000; 2000US-253625P.
 PR 22-DEC-2000; 2000US-257931P.
 PR 16-FEB-2001; 2001US-269308P.
 XX

PA (ELIT-) ELITRA PHARM INC.
 XX

PI Haselbeck R, Ohlsen KL, Zyskind JM, Wall D, Trawick JD, Carr GJ;
 PI Yamamoto RT, Xu HH;
 XX

DR WPI; 2001-611495/70.
 N-PSDB; AAS52261.
 XX

PT New polynucleotides for the identification and development of
 PT antibiotics, comprise sequences of antisense nucleic acids -
 XX
 PS Example 3; Seq ID No 5898; 51pp; English.
 XX

CC The invention relates to antisense inhibitors of genes essential to
 CC prokaryotic cellular proliferation, their use in identifying the
 CC genes, their use in the discovery of novel antibiotics, the essential
 CC genes themselves and the encoded proteins. The prokaryotes used are
 CC *Escherichia coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella*
 CC *pneumoniae*, *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The
 CC invention is also useful for the identification of potential new targets
 CC for antibiotic development. The antisense nucleic acids can also be used
 CC to identify proteins used in proliferation, to express these proteins,
 CC and to obtain antibodies capable of binding to the expressed proteins.
 CC The proteins can be used to screen compounds in rational drug discovery
 CC programmes. The antisense nucleic acid sequence is also useful to screen
 CC for homologous nucleic acids which are required for cell proliferation in
 CC a wide variety of organisms. The present sequence represents an
 CC essential prokaryotic cellular proliferation protein.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 CC
 SQ Sequence 1349 AA;

Query Match 18.7%; Score 576; DB 22; Length 1349;
 Best Local Similarity 30.1%; Pred. No. 5.5e-23;
 Matches 186; Conservative 112; Mismatches 228; Indels 92; Gaps 29;

QY 8 SSDEKNDV---INNOSINTDNNQIKKEETNNYDGIKRSRDEPSTTNVDE-NEAT 63
 Db stklneatsasdgskdvdmqnglndkndngkemsqngnettsngkllleke 114
 QY 64 FLOKTPQDNHILTEEEKSESSVSSNSITDAQPSHTTINRESVQTSNVEDSHVSD 123
 Db svsgltgknkvestaksdeqspkstnedlntkq-----tlsngealq-pdlgenksvvn 168
 QY 124 FANSKIKESNTSGKEENTIEQPNKVKEDSTTSQPSGYTNI-DEKISMODELL--NLPIN 180
 Db -----vgptneenkvd-----aktestt-----lnksdalksndeclvdmnsns 209
 QY 181 EYENKARPLSTSAQP---SIKRVYVNLQALAEQGSVNVNHLIKVTDOSTEGYDSESVI 236
 Db nneennadllpkstapkrlntrmlaavqpssteaknvnlltanltlvddaknmkiv 269
 QY 237 KAHDAENLIYDVFEVDVVKYSGDTMNVVDIDKNV-----PSDLTDSFTIPKIKD-NSG 289
 Db padgylslksqilt--vddkxksqdyftiky-sdlvqyglnpedlkn--lgdlkdpng 323
 QY 290 EIIATGYDNKNKOITTFDYVDKYEENIKAKHLKTSYIDKSKVPNNNTKLDVEYKTALS 349
 Db 382 nttktlanlqypdyvvnknsqsaft--etvshvgnkepfgyyktllynpensnltln 439

```

Cc 324 elatakhdtannllytftdydrfnsvgmglinslymdadtlp--vskndvefnvlig 381
Cc 350 SVNKTIIVEYQRPN--ENRTANLOSMTNIDTKNHT-----VEQTIYNPLRYSAKE 399
Cc 382 ntctktanlgydpdyvneknslgsaft--etvshvgnkempgykqtllynpensln 439
Cc 400 TNVNI-----SGNGDEGSTIIDSTIIKVKYGVGNQNLPSNRRIYDS--EYEDVTNDYA 453
Cc 440 akikvqayhsypnlglnkdvtcliklyqpkgytl---nkgydvntkeltvtn-qy1 495
Cc 454 Q---LGNNDVNINFGNIDSPYIIKVKYDKPNKDDYTTIOQVWTMQTTINEYGEFRTA 510
Cc 496 qkltvgdnnsavldfgnadasyvwmncklfgytinsespcllvqmatlsstgn-----ksv 549
Cc 511 SYDNTIAFSTSSGOGGDLPEPKTYKIGDYVEDVDKDIQNTNDNEKPLSNVLTLYTP 570
Cc 550 stgnalglftnqsgag-----qevykignywedtnknvgel--gekvgvntvltv-fd 602
Cc 571 DGTSSKSVR---TDEDGKY 585
Cc 603 nntnckvgavtkedggy 620

RESULT 8
AAU37544 standard; Protein; 1349 AA.
AAU37544;
AAU37544;
14-FEB-2002. (first entry)
Staphylococcus aureus cellular proliferation protein #1714.
Antisense: prokaryotic cellular proliferation protein;
antibiotic; antibacterial; drug design.
Staphylococcus aureus.
W0200170955-A2.
27-SEP-2001.
21-MAR-2001; 2001WO-US09180.
21-MAR-2000; 2000US-191078P.
23-MAY-2000; 2000US-206848P.
26-MAY-2000; 2000US-207727P.
23-OCT-2000; 2000US-242578P.
23-NOV-2000; 2000US-253625P.
22-DEC-2000; 2000US-257931P.
16-FEB-2001; 2001US-269308P.
(ELIT-) ELITRA PHARM INC.
Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
Yamamoto RT, Xu HH;
WPI: 2001-611495/70.
N-PSDB: AAS55403.
New polynucleotides for the identification and development of
antibiotics, comprise sequences of antisense nucleic acids .
Example 3; Seq ID No 13137; 511pp; English.
The invention relates to antisense inhibitors of genes essential to
prokaryotic cellular proliferation, their use in identifying the
genes, their use in the discovery of novel antibiotics, the essential
genes themselves and the encoded proteins. The prokaryotes used are
Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
invention is also useful for the identification of potential new targets
for antibiotic development. The antisense nucleic acids can also be used

```

```

Cc to identify proteins used in proliferation, to express these proteins,
Cc and to obtain antibodies capable of binding to the expressed proteins.
Cc The proteins can be used to screen compounds in rational drug discovery
Cc programmes. The antisense nucleic acid sequence is also useful to screen
Cc for homologous nucleic acids which are required for cell proliferation in
Cc a wide variety of organisms. The present sequence represents an
Cc essential prokaryotic cellular proliferation protein.
Cc Note: The sequence data for this patent did not form part
Cc of the printed specification, but was obtained in electronic
Cc format directly from WIPO at
Cc ftp.wipo.int/pub/published_pct_sequences.
Cc
Cc Sequence 1349 AA:
Cc
Cc Query Match 18.7%; Score 576; DB 22; Length 1349;
Cc Best Local Similarity 30.1%; Pred. No. 5.5e-23;
Cc Matches 186; Conservative 112; Mismatches 228; Indels 92; Gaps 29;
Cc
Cc 8 SDEBKNDV---INNOSINTDNNQIIRKETNNYDQIERSEDRTSTTNDE-NEAT 63
Cc 55 stnkelnatstaadngsdvdmqngdqndtkndngkemsyqgnatngnkllke 114
Cc 64 FLOKTPQDNTHLTEREEVSSSVSSNSIPTAQQPSHTTINREESVQTSNVEDSHYSD 123
Cc 115 svsgltgknkvestaksdeqaspsctnedlntkg-----ltsngdalq-palgenksvyn 168
Cc 124 FANSKIKESNTESGKEENTIEQPNKVKEDSTTSQPSGYNTI-DEKISQDELL--NLPIN 180
Cc 169 -----vgplneenkkyd-----akestt-----lnvksdal kandelvdnmsns 209
Cc 181 EYENKARPLSTTSQAP----SIKRTVNOQLAEOGSNNVHLIKYTDQISITEGYDSSEVI 236
Cc 210 nmenadlilpkstapklntmrtaavgspteaknvndltntcltvdadknkiv 269
Cc 237 KAHDAENLIYDTFEVDKVKSGDFMTVDIDKNFY-----PSDLTDSFTIPKIKD-NSG 289
Cc 270 pagdylslksqit--vdskvksagdyftlky--sdvqvygllmpedkn---lgdikdpmng 323
Cc 324 elatakhdtannllytftdydrfnsvgmglinslymdadtlp--vskndvefnvlig 381
Cc 350 SVNKTIIVEYQRPN--ENRTANLOSMTNIDTKNHT-----VEQTIYNPLRYSAKE 399
Cc 382 ntctktanlgydpdyvneknslgsaft--etvshvgnkempgykqtllynpensln 439
Cc 400 TNVNI-----SGNGDEGSTIIDSTIIKVKYGVGNQNLPSNRRIYDS--EYEDVTNDYA 453
Cc 440 akikvqayhsypnlglnkdvtcliklyqpkgytl---nkgydvntkeltvtn-qy1 495
Cc 454 Q---LGNNDVNINFGNIDSPYIIKVKYDKPNKDDYTTIOQVWTMQTTINEYGEFRTA 510
Cc 496 qkltvgdnnsavldfgnadasyvwmncklfgytinsespcllvqmatlsstgn-----ksv 549
Cc 511 SYDNTIAFSTSSGOGGDLPEPKTYKIGDYVEDVDKDIQNTNDNEKPLSNVLTLYTP 570
Cc 550 stgnalglftnqsgag-----qevykignywedtnknvgel--gekvgvntvltv-fd 602
Cc 571 DGTSSKSVR---TDEDGKY 585
Cc 603 nntnckvgavtkedggy 620

RESULT 9
AAV83170
AAV83170 standard; Protein; 1802 AA.
AAV83170;
24-JUL-2000 (first entry)
Cell wall protein SdrF.
DE

```



```

XX 21-MAR-2000: 2000US-191078P.
PR 23-MAY-2000: 2000US-206848P.
PR 26-MAY-2000: 2000US-207172P.
PR 23-OCT-2000: 2000US-242578P.
PR 27-NOV-2000: 2000US-253625P.
PR 22-DEC-2000: 2000US-257931P.
PR 16-FEB-2001: 2001US-269308P.
XX
XX (ELIT-) ELITRA PHARM INC.
XX
XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX WPI: 2001-611495/70.
XX N-PSDB: AAS52142.
XX
XX New polynucleotides for the identification and development of
XX antibiotics, comprise sequences of antisense nucleic acids -
XX
XX Example 3; Seq ID No 5779; 511pp; English.

```

```

CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence represents an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp://ipub/pub/published_pct_sequences.
XX
XX Sequence 841 AA:

```

```

Query Match 14.8%; Score 456; DB 22; Length 841;
Best Local Similarity 27.1%; Pred. No. 9e-17;
Matches 158; Conservative 94; Mismatches 222; Indels 108; Gaps 25;

```

```

42 GIEKRSDEPTSTNVDNEATFLOKTPDONTLHTEEVSESSVES---SNSSIDTAQ 97
| : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db 45 gheakaehngelnegkneft-----apsee-----kttkvdvdrqkdhitctad 92
| : : : | : : : | : : : | : : : | : : : | : : : | : : : |
OY 98 QPSHTTINREESVQTSNDVEDSHVDFANSKIKESNTSGKENTLEOPKRVEDSTSQ 157
| : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db 93 gpkvt-----msdatvketssnms-----pqnatangstk 125
| : : : | : : : | : : : | : : : | : : : | : : : | : : : |
OY 158 PSCGTINDEK---ISNDELINLPINYEKARLSTTSQPSK---RVVNOCAA- 208
| : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db 126 tsnvtncksstysneclksnl-----tqakdvstlpttklprtlrmavntvaap 179
| : : : | : : : | : : : | : : : | : : : | : : : | : : : |
OY 209 EOGSNVHLIKVTV--DQSTIEGY-DDESEGVTK--AHDAENLIIVDFEVDKVKSGDPTM 263
| : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db 180 qggltnvdkhfnidialdkghnngtqtefatsvdklklkanytlidssvkegdtft 239
| : : : | : : : | : : : | : : : | : : : | : : : | : : : |
OY 264 VIDDKNTVPSDLUDSTFTPIKIDNGSEIIATGYDNKNKOITTFYVDYKVENIKAHK 323
| : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db 240 fky9qyirpsvrlpsqgtlynaqgnllakgldstntcttlyfnyvdyqytnvsfe 299
| : : : | : : : | : : : | : : : | : : : | : : : | : : : |
OY 324 LTSVIDSKYPPNNNTKIDVEYKTLASSVNKTITVEYGRPNPENTFANQSMFTNIDTKNHT 383
| : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db 300 qvalaktknatldctaykmevtligndyseellvdy---gnkkaqplissstnylnedis 356
| : : : | : : : | : : : | : : : | : : : | : : : | : : : |

```

```

OY 384 VEOFIYIN-PLRYSAKETNV-NISGNGDEGRTIIDSTTIKYYKVGDNQNL-----PDSN 436
| : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db 357 rmtcayvngpkntkytqtlvntt-----gkknpnaknfflyevtdnqvfdsfcpds 411
| : : : | : : : | : : : | : : : | : : : | : : : | : : : |
OY 437 RIYDSEYEDVTNDYAOIGNNNDVINFGNIDS--PYIIKVI-----SKYDPNKDDYTT 489
| : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db 412 klkvdvtdqfdvli---ysndkktatvdlmkqtsnxyllqyavaydnasctdngkldy-- 466
| : : : | : : : | : : : | : : : | : : : | : : : | : : : |
OY 490 IQOFTVMOTITINTEGEFRASIDNTIAPSTSSGOGGLPPEPKTKIKIGYVVEDVDKOG 549
| : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db 467 -----Cltdcktkysw---snysnvngastangd-----qkkylnqdyvvedtnkdg 511
| : : : | : : : | : : : | : : : | : : : | : : : | : : : |
OY 550 IQNTNDNEKPLSNVLTLYTPDGTG--KSVRTDEBGRKQYFDGV 590
| : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db 512 kgda--nekgyikgyvylkdsngkeldrttdtdeangkyqfngl 551
| : : : | : : : | : : : | : : : | : : : | : : : | : : : |

```

```

RESULT 15
AAU37158
ID AAU37158 standard; Protein; 841 AA.
XX
XX AAU37158;
XX
XX 14-FEB-2002 (first entry)
XX
XX Staphylococcus aureus cellular proliferation protein #1328.
XX
XX Antisense; prokaryotic cellular proliferation protein;
XX antibiotic; antibacterial; drug design.
XX
XX Staphylococcus aureus.
XX
XX MO200170955-A2.
XX
XX 27-SEP-2001.
XX
XX 21-MAR-2001: 2001WO-US09180.
XX
XX
XX 21-MAR-2000: 2000US-191078P.
XX
XX 23-MAY-2000: 2000US-206848P.
XX
XX 26-MAY-2000: 2000US-207172P.
XX
XX 23-OCT-2000: 2000US-242578P.
XX
XX 27-NOV-2000: 2000US-253625P.
XX
XX 22-DEC-2000: 2000US-257931P.
XX
XX 16-FEB-2001: 2001US-269308P.
XX
XX (ELIT-) ELITRA PHARM INC.
XX
XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
XX Yamamoto RT, Xu HH;
XX N-PSDB: AAS50117.
XX
XX WPI: 2001-611495/70.
XX
XX New polynucleotides for the identification and development of
XX antibiotics, comprise sequences of antisense nucleic acids -
XX
XX Example 3; Seq ID No 12751; 511pp; English.

```

```

CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence represents an

```

essential prokaryotic cellular proliferation protein.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 CC

Query Match	14.8%	Score 456	DB 22	Length 841
Best Local Similarity	27.1%	Pred. No. 9e-17		
Matches 198	Conservative 94	Mismatches 222	Indels 108	Gaps 25

Tue Jul 30 11:46:38 2002

us-09-147-405-11.ra1

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 30, 2002, 10:28:55 ; Search time 33.65 Seconds

(without alignments)
430.442 Million cell updates/sec

Title: 1
US-09-147-405-11

Perfect score: 3087
Sequence: 1 HHHHHPPSDDEKNDYINN.....SKSVRDEGKQPDGVQVD 593

Scoring table:
BLOSUM62
Gapop 10.0 , Capext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: Issued Patents, AA: *
2: /cgn2_6/ptodata/2/1aa/5A.COMB.pep: *
3: /cgn2_6/ptodata/2/1aa/5B.COMB.pep: *
4: /cgn2_6/ptodata/2/1aa/6A.COMB.pep: *
5: /cgn2_6/ptodata/2/1aa/6B.COMB.pep: *
6: /cgn2_6/ptodata/2/1aa/PCFUS.COMB.pep: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	452	14.6	933	3	US-08-293-728-2
2	452	14.6	933	4	US-09-421-868-2
3	349	11.3	345	4	US-08-856-253-7
4	259.5	8.4	1112	2	US-08-714-402-2
5	259.5	8.4	1161	4	US-09-327-536-2
6	209	6.8	630	4	US-08-973-462-9
7	198	6.4	783	6	5231168-2
8	197	6.4	2314	4	US-09-268-347-49
9	196.5	6.4	1786	4	US-08-973-462-8
10	189.5	6.1	1664	2	US-08-642-846-2
11	189.5	6.1	1664	2	US-09-264-604-2
12	188.5	6.1	1435	2	US-08-568-459A-4
13	188.5	6.1	1435	2	US-08-487-826B-4
14	186.5	6.0	3111	2	US-08-460-309-4
15	186.5	6.0	3111	2	US-08-125-077-4
16	185	6.0	2048	4	US-09-268-347-48
17	184.5	6.0	2391	2	US-08-446-855A-2
18	184.5	6.0	2391	4	US-09-150-741-2
19	177.5	5.7	1087	1	US-08-264-002-5
20	176	5.7	3788	4	US-09-336-447A-76
21	175.5	5.7	1104	4	US-09-268-347-28
22	175.5	5.7	1104	4	US-09-268-347-34
23	175.5	5.7	1183	2	US-08-447-031A-2
24	175	5.7	1147	3	US-08-470-031A-2
25	175	5.7	1147	3	US-08-471-491-5
26	175	5.7	1147	3	US-08-466-662-5
27	175	5.7	3289	2	US-08-477-451-2

28	173.5	5.6	1231	3	US-08-904-263A-4	Sequence
29	173.5	5.6	2123	4	US-08-968-685A-10	Sequence
30	166.5	5.4	984	1	US-08-257-073-3	Sequence
31	166.5	5.4	984	2	US-08-184-009-120	Sequence
32	166.5	5.4	984	2	US-08-458-356-120	Sequence
33	166.5	5.4	984	4	US-08-460-736-120	Sequence
34	166.5	5.4	1164	4	US-08-923-992A-10	Sequence
35	165	5.3	3135	1	US-08-323-110B-2	Sequence
36	165	5.3	3135	4	US-08-954-441-2	Sequence
37	164.5	5.3	1164	4	US-08-923-992A-2	Sequence
38	164.5	5.3	1420	2	US-08-540-804-14	Sequence
39	164.5	5.3	1420	2	US-08-218-265-14	Sequence
40	164.5	5.3	1420	3	US-08-521-872-14	Sequence
41	164.5	5.3	1420	4	US-08-590-399-14	Sequence
42	163.5	5.3	1098	4	US-08-923-992A-8	Sequence
43	163.5	5.3	1128	4	US-08-923-992A-6	Sequence
44	162.5	5.3	740	1	US-08-257-073-5	Sequence
45	162.5	5.3	1964	2	US-08-790-912-3	Sequence

ALIGNMENTS

RESULT 1
US-08-293-728-2
Sequence 2, Application US/08293728D
Patent No. 6008341
GENERAL INFORMATION:
APPLICANT: Foster, Timothy J.
TITLE OF INVENTION: The S. aureus Fibrinogen Binding Protein Gene
FILE REFERENCE: 05344.105011
CURRENT APPLICATION NUMBER: US/08/293,728D
CURRENT FILING DATE: 1994-08-22
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 933
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-08-293-728-2

Query Match	14.6%	Score 452;	DB 3;	Length 933;
Best Local Similarity	26.6%	Pred. No. 1.6e+18;		
Matches 154;	Conservative 105;	Mismatches 234;	Indels 86;	Gaps 26;
QY	8	SSDEKNDYINNOSINTDDNOI	IKKEETNNYDGIKRS	EDRTESTYVDENEATFLQK 67
DB	51	SNESKSDSSVSAARKTDITNV	-----SDKRTSSNTN	---NGETSAVO 91
QY	68	TPQDNTHLTEBEYKSSSVSSNS	IDTQAQTSHTTINNEES	VQTSVDNEDSVDFANS 127
DB	92	NP-----AQOETVOSSSNAT	TEETPVGEATTTTNOAN	PATQOS-SMTNAEELVN- 143
QY	128	KIKESNTESGKEENTTEQPNK	VEDSTGOPSQGTNIDEK	ISNOELNLN-LPINEYENKA 186
DB	144	--QTSSETTFNDITNV-----	SSVNSPONSSTAENVST	QDITSTPATBN--NES 189
QY	187	RPLSTSAOPSIKRYVN-----	OLAAEGSNVNLIKYTDQ	-SITEGYDSEGV 235
DB	190	APQSTASAKNDVYNQAVN	SAPRBAFSLAAVAAD	PAAGTDITNLQTNVYGI -DSGTT 248
QY	236	IKAHDAENLIYDVEYDKV	SGDTMYDIDKNYPSDIT	SDTSFTPKIKNSGELIANG 295
DB	249	VYFHQAGYKLVNGFVPS	SAVAGDTEFKITVPEK	ELNNGVSTAKVPPIAGD-QYLANG 307
QY	296	TYDNKKKQITYPEPTD	VDVDEYENIKAKHLK	LSYIDKSKVNNNNKL-DVEYKTALS--VN 352
DB	308	VIDS-DGNVITYFTD	IVNKKDKVAKLTMP	AYTD---PENVAKTGNVLAIGSGTTAN 362
QY	353	KITVEYQRPENRNTAN	LSMFTNIDTKNHTVE	QTIYINPL--RYSAKETNNVNSGQDE 410

Db 363 KTVLVDEYKGFYNLISIKGIDIDQIDKTNNTYQRIIVYVNSGDNVIAPIVLGNLKRNTDS 422
QY 411 GSTIIDDSTIKYKKGDNONLPSNRITDYSEYEDVTDVAQLGNNDVNIENG- 466
Db 423 NALIDQONTSIKRYKVDNADLSESYFV-NPENFEVDYNSVNTFFPNPQYKEEFTPPD 481
QY 467 NIDSPYIIVISKYDPN-KDDYTTIOQYTMOTLINEYGE--FRTASDNTIARSTSSG 523
Db 482 QITPTIIVYVNGHIDNSGID-----LALRSTLYGYSNIIWMSMDNEVAFNNGSG 534
QY 524 OGOG-DLP--PEKTYKIGDY--VMEVDVKD-GIONTNDN 556
Db 535 SCBGIDKPYVPEQDPDEGEIEPIEDSDSDPSGSDS 573

RESULT 2

US-09-421-868-2
Sequence 2, Application US/09421868
Patent No. 6177084
GENERAL INFORMATION:
APPLICANT: Foster, Timothy J.
APPLICANT: McDevitt, Damien L.
TITLE OF INVENTION: The S. aureus Fibrinogen Binding Protein Gene
FILE REFERENCE: 05344.105011
CURRENT APPLICATION NUMBER: US/09/421,868
CURRENT FILING DATE: 1999-10-19
PRIOR APPLICATION NUMBER: 08/293,728
PRIOR FILING DATE: 1994-08-22
NUMBER OF SEQ ID NOS: 20
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 2
LENGTH: 933
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-09-421-868-2

Query Match

Best Local Similarity 14.6%; Score 452; DB 4; Length 933;
Matches 154; Conservative 105; Mismatches 234; Indels 86; Gaps 26;

QY 8 SSDEKNDVYNNQSDINTDNNQIKKEETNNYDIEKRSDEKSTSTVNDENAEFLQK 67
Db 51 SSSKSNDSVSAAPKTDITV-----SDPKTSNTN--NGERSVQ 91
QY 68 TPQDNHLEEEKSESSSSSIDTQAQPSHTTINEESVQSDNVEDSHVDFANS 127
Db 92 NP-----AQETQSSSTNATTEETPTVGTGATTTTQANPTATQOS-SMNAELVN- 143
QY 128 KIKESNTESGKEENTIEQPNKYKEDSTVQPSGYTNIIDEKISNOBELN-LPINEENKA 186
Db 144 --QTSNETFNDITV-----SSVSPONSTNAEVSSTODTSTATISN--NDS 189
QY 187 RPLSTTSAOPSIKRYTVN-----QLAEOGSNVNHLIVTQD--SITEGYDSECV 235
Db 190 APOSTDASKNDVANOAVNTSAPRMRAFSIAVAADAPAGCTDITNQLTNVTGI-DSCTT 248
QY 236 IKAHAENLIYDVTPEVDKVKSGDMYDIDKNTVPSDLTFSITPKKDSGELIATG 295
Db 249 VTPHQAGYKLVGFSVPNSAVKGDFTKIVPEKELNLCVSTAKVPPIMAGD-QVLANG 307
QY 296 TYDNKKQITVTFDQVDEYENIKAHKLTSYIDSKVYNNNTKL-DVEYKTALSS--VN 352
Db 308 VIDS--DGNVIYTFDQVNNKDKVAKILMPAYID----PENKKGAVNLATGIGSTYAN 362
QY 353 KTIIVYQRPENRFTANLQSMFTNIDTKNHYEQTIYINPL--RYSAKETNVNIGNGDE 410
Db 363 KTVLVDEYKGFYNLISIKGIDIDKTNNTYQRIIVYVNSGDNVIAPIVLGNLKRNTDS 422
QY 411 GSTIIDDSTIKYKKGDNONLPSNRITDYSEYEDVTDVAQLGNNDVNIENG- 466
Db 423 NALIDQONTSIKRYKVDNADLSESYFV-NPENFEVDYNSVNTFFPNPQYKEEFTPPD 481

QY 467 NIDSPYIIVISKYDPN-KDDYTTIOQYTMOTLINEYGE--FRTASDNTIARSTSSG 523
Db 482 QITPTIIVYVNGHIDNSGID-----LALRSTLYGYSNIIWMSMDNEVAFNNGSG 534
QY 524 OGOG-DLP--PEKTYKIGDY--VMEVDVKD-GIONTNDN 556
Db 535 SCBGIDKPYVPEQDPDEGEIEPIEDSDSDPSGSDS 573

RESULT 3

US-08-856-253-7
Sequence 7, Application US/08856253
Patent No. 6288214
GENERAL INFORMATION:
APPLICANT: Hook, Magnus
APPLICANT: Patil, Joseph M.
APPLICANT: House-Pompeo, Karen
APPLICANT: Schanm, Narayana
APPLICANT: Symersky, Jindrich
TITLE OF INVENTION: COLLAGEN BINDING PROTEIN COMPOSITIONS
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: U.S.
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/856,253
FILING DATE: Concurrently Herewith
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/017,678
FILING DATE: 16-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: Kitchell, Barbara S.
REGISTRATION NUMBER: 33,928
REFERENCE/DOCKET NUMBER: TAWK:193
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (512) 474-7577
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 345 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-856-253-7

Query Match

Best Local Similarity 11.3%; Score 349; DB 4; Length 345;
Matches 102; Conservative 61; Mismatches 136; Indels 32; Gaps 15;

QY 220 VTDQ--SITEGVDSEGVYKHAENLIYDVTPEVDKVKSGDMYDIDKNTVPSDLTD 277
Db 25 ITNQLTNVTGI-DSGTVTPHQAGYKLVGFSVPNSAVKGDFTKIVPEKELNLCVST 83
QY 278 SFTIPKINDSGEILATGYDNKQKITVTFDQVDEYENIKAHKLTSYIDSKVFPNNN 337
Db 84 TAKVPPIMAGD-QVLANGIDS--DGNVIYTFDQVNNKDKVAKILMPAYID----PENV 137
QY 338 TKL-DVEYKTALSS--VKKTIIVYQRPENRFTANLQSMFTNIDTKNHYEQTIYINPL- 393
Db 138 KKTGNVTLATGIGSTANKTVLVDEYKGFYNLISIKGIDIDKTNNTYQRIIVYVNSG 197

QY 394 -RYSAKETNNVNSGNGDEGSIITIDSTIIKYYKGDNONLPD3NRIYDSEEDYNDY 452
Db 198 DNVIAPVLGTGMLKRPDTSNALIDQNTSIKYYKDNADLSESYFV-NPENEEDYTSVN 256
QY 453 AQLGNNDVNINFG---NIDSPYIIKYSKYDPN-KDYTTTIOQVMTQTTINEYGE- 506
Db 257 ITFPNPQYKVEFNTPPDQITTPYIVVYNGHIDPRSKD-----LALRSLIYGNINI 309
QY 507 -FRTASYDNITAFSTSGGOG-DLP--PEK 533
Db 310 IWRSMNDNEVAFNNGSGSGDIDKPYVPEQ 340
RESULT 4
US-08-714-402-2
Sequence 2, Application US/08714402
Patent No. 5910441
GENERAL INFORMATION:
APPLICANT: ROCHA, Claudia
TITLE OF INVENTION: FIBRONECTIN AND FIBRINOGEN BINDING
TITLE OF INVENTION: PROTEIN FROM GROUP A STREPTOCOCCI
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: BURNS, DOANE, SNECKER & MATHEIS
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/714,402
FILING DATE: 16-SEP-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 016921-097
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1112 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-714-402-2
Query Match 8.4%; Score 259.5; DB 2; Length 1112;
Best Local Similarity 21.2%; Pred. No. 2.3e-07;
Matches 116; Conservative 102; Mismatches 205; Indels 123; Gaps 23;
QY 132 SNESGKEENTIEOPKVKEDSTTSOPS-----GYTINDEKISNODEL--NLPIN 180
Db 43 ANKQGFELIKKYDONKKPLPGATSSLSLTKDGKTSVOSTSNDKGIYDAQNLQPGTYTLK 102
QY 181 E-----YENKARPLSTSAQPSIKRYTVN---QLAAEGS-NVNHLLIKVTD--OSTIE 227
Db 103 EEARPGYDKTSRTWTVYVYENGTYKLVENPYNGEIIKSKAGSVSSQLGLEPKMSVYS 162
QY 228 GYDDSE---GVIAHAENLIYVTEFVDDKYS-----GDPTVVIDIKNTVPSDLTDSF 279
Db 163 KYGKTEVSSGAADFYRNHAAVYFKMSFELKOKDKSETINPGDTFVLQIDRLRLNPKGISQD- 221
QY 280 TIRK-INKNGEIIATGTYNKKKQITVTFDVIDKYENIKAHILKLSYIDKSKVPPNNT 338

Db 222 -IPKIIIDSAISPPLAIGYHNEHNLITFTFDYIAGLDKVLASLSELENKTELEN- 278
QY 339 KLDVEYKTALSSVYKTI-----TVEYORPNE-----NRTANLOSMTNIDTK 380
Db 279 -----TSISNFKSTIGGEITYKGTAVNLVYGNESKTESNYITNGLSNMGSIESTYTE 331
QY 381 NHTVEQITVYINPLKYSAKETVNVISNG-----DEGSIITIDSTIIKYYKVGNDQ 430
Db 332 TGEFVWYVYVNPNTNIPYATNMILGFEGRANSNTSDELDNANTSAGELIQVLEVEPGE 391
QY 431 NLPSDNRIYDSEYEDVT-----NDGYAQLGNNDVY---NINGN--IDSPYIIKYSK 479
Db 392 KLPSSYGV-----DYKLLRLRDITAGLNGFGMTKRQRIIDGNINQNAFIITKYVK 444
QY 480 YDPNKDYTTIOQVMTQTTINEYTGERTAS-----YDNTIAFSTSGGOG--- 527
Db 445 TDOS-----GKPLVQSNLASFEGASEYAAFTPVGNGVYFONEIALSPKSGSGSKSE 497
QY 528 ----DLPEKTYKIGDYWEDVDKGIQNTNDNEKPLSNVLTLYTPDGTSKSVR--TDE 581
Db 498 FTKPSITVANKRAVLAQLRKKMTSDNV-----PLPEAAFEIILSSNGNSOKLEASSVT 549
QY 582 DKYQF 587
Db 550 QGEVHF 555
RESULT 5
US-09-327-536-2
Sequence 2, Application US/09327536
Patent No. 6355477
GENERAL INFORMATION:
APPLICANT: FISCHETTI, Vincent A.
TITLE OF INVENTION: FIBRONECTIN AND FIBRINOGEN BINDING PROTEIN FROM GROUP A
TITLE OF INVENTION: STREPTOCOCCI
FILE REFERENCE: 022927-008
CURRENT APPLICATION NUMBER: US/09/327,536
CURRENT FILING DATE: 1999-06-08
PRIOR APPLICATION NUMBER: US 08/714,402
NUMBER OF SEQ ID NOS: 2
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 2
LENGTH: 1161
TYPE: PRT
ORGANISM: SFFBP gene
US-09-327-536-2
Query Match 8.4%; Score 259.5; DB 4; Length 1161;
Best Local Similarity 21.2%; Pred. No. 2.4e-07;
Matches 116; Conservative 102; Mismatches 205; Indels 123; Gaps 23;
QY 132 SNESGKEENTIEOPKVKEDSTTSOPS-----GYTINDEKISNODEL--NLPIN 180
Db 43 ANKQGFELIKKYDONKKPLPGATSSLSLTKDGKTSVOSTSNDKGIYDAQNLQPGTYTLK 102
QY 181 E-----YENKARPLSTSAQPSIKRYTVN---QLAAEGS-NVNHLLIKVTD--OSTIE 227
Db 103 EEARPGYDKTSRTWTVYVYENGTYKLVENPYNGEIIKSKAGSVSSQLGLEPKMSVYS 162
QY 228 GYDDSE---GVIAHAENLIYVTEFVDDKYS-----GDPTVVIDIKNTVPSDLTDSF 279
Db 163 KYGKTEVSSGAADFYRNHAAVYFKMSFELKOKDKSETINPGDTFVLQIDRLRLNPKGISQD- 221
QY 280 TIRK-INKNGEIIATGTYNKKKQITVTFDVIDKYENIKAHILKLSYIDKSKVPPNNT 338
Db 222 -IPKIIIDSAISPPLAIGYHNEHNLITFTFDYIAGLDKVLASLSELENKTELEN- 278
QY 339 KLDVEYKTALSSVYKTI-----TVEYORPNE-----NRTANLOSMTNIDTK 380

Db 279 -----TSISNFKSTIGGQETTYKGVNVLGNSTKESNTITNGLSNVGSIESYNT 331
QY 381 NHVEQTIYNPLRYAKETNNVINGNG-----DEGSTIIDSTTIKVKYGDNO 430
Db 332 TGEFVYVYVNPNTNIPYATNMLGFRARSNTSDLENDANTSSELGEIQYVEPEGE 391
QY 431 NLPDSNRIDYSEYEDVT-----NDYAOGLNNNDY-----NINEGN--IDSPYIIVISK 479
Db 392 KLPSSYGV-----DVTKLITLTDITAGLGNFQMTKQORIDFGNNIQKAFIIVTKG 444
QY 480 YDPNKADYTIQOTYMTQTTINEYTGEPRTAS-----YDNTIAFSTSGOGOG--- 527
Db 445 TDS-----GKPLVVOXNLASFSGASEAAFTPVGVNVRQNEIALSPKSGSGSKSE 497
QY 528 ----DLPEKTYKIGDYVEDVDKQIQNTNDNEKLSLVLTLPDGTSKSVR--TDE 581
Db 498 FTKPSTTVANLKRVAOLRKKMSTDNV-----PLPEAFELRSSNGNSOKLEASSNT 549
QY 582 DKRYOP 587
550 QGEVHF 555

RESULT 6
US-08-973-462-9
; Sequence 9, Application US/08973462B
; Patent No. 6191270
; GENERAL INFORMATION:
; APPLICANT: DRUILHE, PIERRE
; APPLICANT: DAUBERSIES, PIERRE
; TITLE OF INVENTION: MALARIAL PRE-ERYTHROCYTIC STAGE POLYPEPTIDE MOLECULES
; FILE REFERENCE: 0660-0125-0 PCT
; CURRENT APPLICATION NUMBER: US/08/973,462B
; EARLIER FILING DATE: 1998-02-06
; EARLIER APPLICATION NUMBER: PCT/FR96/00894
; EARLIER FILING DATE: 1996-06-12
; EARLIER APPLICATION NUMBER: FR 95/07007
; EARLIER FILING DATE: 1995-06-13
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 9
; LENGTH: 630
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Polypeptide
US-08-973-462-9

Query Match 6.8%; Score 209; DB 4; Length 630;
Best Local Similarity 19.5%; Pred. No. 8.6e-05;
Matches 116; Conservative 121; Mismatches 206; Indels 152; Gaps 28;

QY 9 SDEKNDVYNN-----NOSINTDNNQIIEKETNNYDGIEKSEDETEST- 54
Db 59 SDELNELNLNDVNGVEKENLIESQVNDIFNLSLVQOEOOHNEEKEVESEEND 118
QY 55 -----TJVDENE-----ATFLCK-----TQODNHLLEEVK 81
Db 119 EESVEENVEENVEENDSDSSVASSVESIDSSIEINAPVTEELIAPVTEIV 178
QY 82 ESSSVSSSSSIDTAQPSHTTINREESVQTSNDVEDS---HVSDFANSKIKES---NTE 135
Db 179 APSVSEVAAPVSEESVARS-----VEESV--AENVEESVAVNEVEIAPVSEESVANEVE 231
QY 136 SKGKENTIEQNKYKEDSTTQPSGTYNIDKIS--MODELINLPIN-----YENKARPL 189
Db 232 ESVAENVEESVAVNEVEESVAV-----NVEESVAVNEVEIAPVTEESVAPVTEIAPV 285
QY 190 STTSAPSKIRVTYNOLAEOGSNNHLIKVTDSITGCVYDSDSGVIAHAEMLIDYT 249
Db 286 VEESEAPVTEIAPVSEESVAPSV-----EESVAVNEVEESV-----AENVEESVAV 331

QY 250 FEYDVKYKSG--DTMTVIDDKNTVPSDLNDSTFIPIKINDNSEIATGYDNKKNOI--- 304
Db 332 ENVEESVAVNEVEESVAVNEVEIAPV--VEEIIAPVSEESVAVNEVATNLSNLSLLG 389
QY 305 --TYFTDVV--KYEIKAKHLKLTSTIDSKYPNNNTKIDVEKTLSSV----- 351
Db 390 IETEEIKDSILNEIEVEKENV--VTILKEVE-----ETIASVYTFSLIEIOENTIT 442
QY 352 NRTIYVEXORPNEPNTANLQSMFTNIDTKNHVEQTIYNPLRYAKETNNVINGNGDGC 411
Db 443 NDTIEKLELHEN--VLSALENQSEKKEVIDVEEVEKAVATLIEVEAEAE 499
QY 412 STIIDSTTIKXY-----KVGON--ONLPDS--NRIYDSEY-----EDVND 450
Db 500 S-----ESTITTEIFENLEENAVESNEKVAENLEKNETYFNTLDKYEETVEISGESLENN 555
QY 451 D-----YAOGLNN-----NDVNINGNDSPYIIKIVSKYOPNKDDYTTIOQV 494
Db 556 EMDKAFSEIFDYNVKGIOENLITGMFRSITISVIOSEEKVDLNEVSSIIDNI 610

RESULT 7
5231168-2
; Patent No. 5231168
; APPLICANT: DIEGIEL, MORTEN; BORRE, MARTIN; JEPSEN, SOREN;
; VUUST, JENS; RIENECK, KLAUS; WIND, ANNETTE; JAKOBSEN, PALLE H.
; TITLE OF INVENTION: MALARIA ANTIGEN
; NUMBER OF SEQUENCES: 19
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/409,658
; FILING DATE: 18-SEP-1989
; SEQ ID NO: 2
; LENGTH: 783
5231168-2

Query Match 6.4%; Score 198; DB 6; Length 783;
Best Local Similarity 18.5%; Pred. No. 0.00049;
Matches 118; Conservative 127; Mismatches 216; Indels 176; Gaps 28;

QY 4 HHHPSDEKNDVYNNNSINTD-----DNNQIIEK-----EETNNYDG 42
Db 54 HHENSSSEVSNLENENFEVSDKVTPEAEHEVEESSESNPEAPNESSIEEHOEII 113
QY 43 IEKSEDETESTYNDENEATFLCKT-----PQDNTLIEEVKSSSVSSSSSIDTAQ 97
Db 114 VPEQNDSESGGLVDNEEGDFEEPNHFEFPDQ---DSBLSENLIEVESKVSSEPAE 169
QY 98 QPSHTTINRESV--QTSNDVEDSHVSDPANSKIKESNTESKRENTIEQPKVKEDSTS 156
Db 170 ---HVEIYSEKVSSEPAHVE--IVSEKSTSEPAE--HVESVSEDSN--NEPSEKKGVPVS 222
QY 157 QPSGTYNIDKISMODELINLPINEYE--NKARPLSTTSAPSKIRVTYNOLAEOGSNNV 215
Db 223 KP--FEEL--EKVDQPKYVDLQIIEPNVVDQPNQPEVEPSF--VKEKPESEKNKIAS 277
QY 216 HLIVTDSITGIDSDSGVIAHAEMLIYDVEFVDVKYSGDTMTVIDDKNTVPSDL 275
Db 278 -----VPEVEKEKEVSEVEEKOQSVSEIPIVN--EDEREVDHTQDLDDHKTVDPEI 331
QY 276 TDSFTPIKIKNSGEIATGTYYDNKKNOITFTFDYVDKIENIAHLKLTISYIDSKYVP 335
Db 332 VEVEIIS-----ELHENEVAHVEIIVEIEVEFEPN 362
QY 336 NNTKLIDVEYKTALSSVKITVEYQRPNE--NRTANLQSMFTNI-----DTRKNHTVEQTI 388
Db 363 QNN-----EFQIENEDKSKAHIGHIEIYVEEILPEDDKNEKVEHEI 403
QY 389 YINPLRYAKETNNVINGNGDSTIIDSTTIIVYKGVNQLPDSNR---IYDSEY 444
Db 404 -----VEVEIILPEDKNEKGOHEIIVEEII--LPEDDKNEKVEHEIYEV 445
QY 445 EDVTNDYAOGLNNNDVNNINGNDSPYIIKIVSKYOPNKDDYTTIOQVYMTQTTINEYT 504

DB 446 EELPEDKNEKGOHEIVE-----EILPEDKNEKEHEIVEEILPEDKNE-- 493
QY 505 GEFRRASYNTIAFTSSQOGOD-----LPRETTYIGDYWEDVDKDGQNTNDNE 557
DB 494 -----KGOHEIVEEELPEDKNEVOH--EIVEEELPEDKNE 531
QY 558 KPLSNVL-VLTATPDGTSKSVRTDEDKYQFDQVQD 593
DB 532 KGOHEIVEEELPE-----EDKNEKGOHEIVE 561
RESULT 8
US-09-268-347-49
Sequence 49, Application US/09268347
Patent No. 6335182
GENERAL INFORMATION:
APPLICANT: Loosmore, Sheena M.
TITLE OF INVENTION: RECOMBINANT HAEMOPHILUS INFLUENZAE ADHESIN PROTEINS
FILE REFERENCE: 1038-860
CURRENT FILING DATE: 1999-03-16
CURRENT FILING DATE: 1999-03-16
NUMBER OF SEQ ID NOS: 54
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 49
LENGTH: 2314
TYPE: PRP
ORGANISM: Haemophilus influenzae
US-09-268-347-49

Query Match 6.4%; Score 197; DB 4; Length 2314;
Best Local Similarity 21.1%; Pred. No. 0.0022;
Matches 135; Conservative 101; Mismatches 261; Indels 144; Gaps 30;

QY 11 EEKNVYINNOSIN-TDNNQIIKKEETNNYDIEKRSDETESTVNDENATFLQKTP 69
DB 379 EFGADKLTPNNNNIGVYTDNNTGLKAKLNSGLE-----TVSTKMLTASE---KTVV 428
QY 70 QDTHLHEEBEVESSSVSSSIDTAQPSHTTINREESVQSDN---VEDSHVSDFA 125
DB 429 GSGNNNAELQSGGLTPTTMASTD-----KTVGTDLKFTDSDNVALEDT--TRIT 479
QY 126 NSKIKESNTESGRENTIEOP---NKVKEDSTTSOPSGYTINDEKISNODELINLPINE 181
DB 480 KOKIIGSNKAGTYDEN---KPYLDKOKLKGNTLNGGLT-VNNTIGGSNKQIQV---- 531
QY 182 YENKARPLSTSAQPSIKRYTVNQLAEOGSNVNHLIKVTDOSTIE--GYDSEGVYKA 238
DB 532 -----GADGIKFADVN-----VNVYNAKFGTTRITIEBEIGFADADGKYDK 572
QY 239 H-----DAENL-IYDYTFEVDKYSQDTPMVDIDKNTVPBDLDSFTIPIK----DNNG 289
DB 573 KSPYLKQKQOVGVKITKQSGINAGOKISNVKDAT---DDTDVAVYKQDKQYQDADG 629
QY 290 EIIATGYDNKKQIT---YTFDYDYKYENI---KAHLKLISYDKSKVPNNNTKLD 341
DB 630 ALOSFTIRDEKGOEFTISNLXSNQNTPEFTIFAGEINGISINDAKGV-----KVG 684
QY 342 VEKKTALISVNAKTYVEYQRPENRANLOSMTNIDTKHHTVEQTYIINPLRSAKETN 401
DB 685 IDPLNGLTGKPLV-----GSDKDGKTQLVLEOVASGCTKNCILINGVSPLEPSI--TN 735
QY 402 VNISGNDESGTIIDSTIIKVKYVGNQNL-----PDSNRIVYXSEYEDVTNDYQOLG 456
DB 736 AGGVRTTEQNTTSDCKRAASIGILNTGENLKNNSVGVSTYNTVDFID-----G 791
QY 457 NNDVNINFGNISPYIIRIKYISKYDPRKDDYTT-----IQQVTMQTTINEY 503
DB 792 NATAKATYDETQNTSKV-----TYDVAVDEKTEILLDNGKNTNKIGVKTTLTTLTNANGK 847
QY 504 TGEFRNASYD-----NTIAFTSSQOGODLPREKTYKI-GDYVWED-----VD 546

DB 848 ATNFSTTDNALVNAKDIAENINFLAKEIHTTKGADTALQ-TFKYKKGATDDEFTYVG 906
QY 547 KDGIQNTDNEKPLSNVLVLTATPDGTSKSVRTDEDKYQF 587
DB 907 KQGTQNT-----GKVTNTLKLKGNGLTVAATNKGTYTF 939
RESULT 9
US-08-973-462-8
Sequence 8, Application US/08973462B
Patent No. 6191270
GENERAL INFORMATION:
APPLICANT: DRUILHE, PIERRE
APPLICANT: DAUBERSIES, PIERRE
TITLE OF INVENTION: MALARIAL PRE-ERYTHROCYTIC STAGE POLYPEPTIDE MOLECULES
FILE REFERENCE: 0660-0125-0 PCT
CURRENT APPLICATION NUMBER: US/08/973,462B
CURRENT FILING DATE: 1998-02-06
EARLIER APPLICATION NUMBER: PCT/FR96/00894
EARLIER FILING DATE: 1996-06-12
EARLIER APPLICATION NUMBER: FR 95/07007
EARLIER FILING DATE: 1995-06-13
NUMBER OF SEQ ID NOS: 29
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 8
LENGTH: 1786
TYPE: PRP
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Polypeptide
US-08-973-462-8

Query Match 6.4%; Score 196.5; DB 4; Length 1786;
Best Local Similarity 18.7%; Pred. No. 0.0017;
Matches 128; Conservative 142; Mismatches 264; Indels 149; Gaps 26;

QY 21 QSINTDNNQIIKKEETNNYDIEKRSDETESTVNDENATFLQKTPDQNTHLTEEV 80
DB 549 ESVEENVEESAENVEESAENVEESAENVEE--SVAENVEEIVAPVEEIVAPVEE 606
QY 81 KESSVSSESSNSIDTAQPSHTTINREESVQSDNVEDSHVSDFANSKIKESNTESGKE- 139
DB 607 VAPSVESVAPSVESVEE-----NVEESV--AENVEES-----VAENVEESAENVEE 654
QY 140 -----ENTIEQ---PKVKEDSTTSOPSGYTINDEKIS-NODELINLPINEYENKA 186
DB 655 VAPVEEIVAPVEEIVAPSVESVAPSVESVAPSVESVAPSVESVAPSVESVAPSV 712
QY 187 RPLSTTSNQ-----PSIKRV---TVNQLAEOGSNVNHLIKVTDOSTIEGYDSEGVY 236
DB 713 ENVEESAENVEEIVAPVEEIVAPVEEIVAP--SVESVAPSVESVAPSVESVAPSV 767
QY 237 KAHDAENLIDYTFEVDKYSQ--DTMYVDIDKNTVPBDLDSFTIPIKIDNSGELIAT 294
DB 768 -----AENVEESAENVEESAENVEESAENVEESAENVEESAENVEESAENVE 821
QY 295 GTYDNKKQI-----YTFDYDY-DKYENIKAHKLISYDKSKVPNNNTKLDVEYKTA 348
DB 822 NLSQNLISNLGIEIEEIDISINIEEYKENVYTT-----ILEVETTAESVYTF 874
QY 349 SSV-----NKITVEYQRPENR-RTANLOSMTNIDTRN----- 381
DB 875 SNLEEIQENTTINDTIEKLELHEHVNLSALENTGSEKKEKVEIDIVEKVEEAVATL 934
QY 382 HTVEQ-----TYINPLRSAKETVNVNISGNDESGTIIDSTIIKVKYVGNQNL 431
DB 935 IEVEQAEEKSANTITEIFENLEENAVESNENNAENLEKNEVEYFNVLDKVEEVEISG 994
QY 432 LPSNRIRYDSEVEDVNDYQOLGNNDVNINNGIDSPYIIRIKYISKYDPRKDDYTT- 490
DB 995 ESELENEMDKAFPSF--DNVKGIOENLTLGAFRSIETISIVIOSEKVEVLNENYVSSIL 1052

QY 491 -----OQVYMTOTINEXYGERFTASNDTIAFTSSG----- 523
DB 1053 DNENKKEGLINKLENISSTEGVETVEHEQNVYVDVFPAMKQOFLGLNEAGLKE 1112
QY 524 -----OGGDDLPEPKYKIGDYWEDVDKQIONDNEKPLSNV-----VTLT 568
DB 1113 MFENLEDFVSESDVITVEIK-DEPVQKREKEVETIIEKEMENYDVLEEKEDULTK 1171
QY 569 YPDGTSKSVRTDEGKYQDFDQV 591
DB 1172 MDAVESIEISSDSKEETESIK 1194
RESULT 10
US-08-642-846-2
Sequence 2, Application US/08642846
Patent No. 5886151
GENERAL INFORMATION:
APPLICANT: HOSTETTER, MARGARET K.
APPLICANT: GALE, CHERYL A.
APPLICANT: BENDEL, CATHERINE M.
APPLICANT: TAO, NIAN-JUN
APPLICANT: KENDRICK, KATHLEEN
TITLE OF INVENTION: CANDIDA ALBICANS GENE, INTEGRIN-LIKE
TITLE OF INVENTION: PROTEIN, ANTIBODIES, AND METHODS OF USE
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESSES:
ADDRESSEE: MUEITING, RAASCH, GEBHARDT & SCHWAPPACH, P.A.
STREET: 119 NORTH FOURTH STREET, SUITE 203
CITY: MINNEAPOLIS
STATE: MINNESOTA
COUNTRY: USA
ZIP: 55401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/642,846
FILING DATE: 03-MAY-1996
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: MUEITING, ANN M.
REGISTRATION NUMBER: 33,977
REFERENCE/DOCKET NUMBER: 110.00280101
TELECOMMUNICATION INFORMATION:
TELEPHONE: 612-305-1217
TELEFAX: 612-305-1228
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1664 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-642-846-2
Query Match 6.1%; Score 189.5; DB 2; Length 1664;
Best Local Similarity 20.4%; Pred. No. 0.0039;
Matches 131; Conservative 105; Mismatches 236; Indels 171; Gaps 31;
QY 15 DVINNOS---INTDNOIIRK-----BETNNOGIEKRSRDREEST 55
DB 138 DKVENHAPFYINTSPKSKTKATPKASKKAAFTVTNEIHIHP--DNKVEEDDSQ 135
QY 56 NVDENATFLOKTPQDNTHLTEEVKSSSVSSNSIDTAQOFSHTT-----INRE 108
DB 196 KEDSVEPPLQ-----HQMKDPSQFNYSDEDTNASVPT-PLKHTKTPAQLLNKN 247
QY 109 SVQTSNVEVSHSD-----FANSKIKE-----SNTSGKEENTIEDPNKVKEDST 135

DB 248 EV-----NSREPALTDKMLKRENFNSLDEKVALYLSPTNNNNKNSVSDMSHLQ-----N 299
QY 156 SPSGYTINDEKISNODELLNPINEYENKARPL-STSAQPSIKRYTVNQLAEQGSVY 214
DB 300 LODASKNKTNNENHNSFALKAPKNDIEN---PLNSLTYNADISLRSSGSSOSLSLRND 356
QY 215 NMLIKV---TDOSITGYDSDSEGVIRAKHAENLIYVTEFVDKVKSGTMTYDIDKNY 271
DB 357 NRYLEVPSPSKVNPGLSLNDIGKFS-----EVEESLT 392
QY 272 PSDLT-DSEFTIPKID-----NSGEIANGTYDNKKNQIYTFYDVKYENIKAH----- 321
DB 393 PRDLSRDKLETTKEHDAPENHNNENFIDAKSTYNNKG-QLYVSSDHLIDSDRSYNNTEGS 451
QY 322 -LKIISYIDSKVYNN-----NTLIDVEYKTAALSSVKN 353
DB 452 ILNLNLSAQSQISLNALEKOROTOEOQAPEPEETSFSNINIKVQKPKNLEFVYV 511
QY 354 TIVEYQRENENTANLQSMFTNIDTKNH-TVEQTIYINPLRYSAKETVNNISG----- 406
DB 512 TIKEPVSAATEIKAPKREFSSRLRIKNEDEINAPADIHPPKKNENANSHVEDTALLKRA 571
QY 407 -NGDEGSIIDSTIIRV-YKVGDNQNPDSNRIRYDSEVEDYTNDDYQAGNN---NDV 461
DB 572 LNDDESPTQNSTKMSIRPHDSQKLEDSND---GRED--NDQISRFKSDILNDV 625
QY 462 N-----INFGNIDSPYIIKYSKIDPNKDYTTIQOYVYMTOTINEXYGERF-----T 509
DB 626 SOTSIDIIGKYGSSSEIITKTJLA---PPRSDNNDXENSKSLDPPANNESLQOOLEVPH 682
QY 510 ASYDNTIARFSTSGQOGDLPPEK---TYKIDGY-VWEDVDK 547
DB 683 KEDDSILANSSNIA-----PPEELTPVYEANDYSSFNDYTK 719
RESULT 11
US-09-264-604-2
Sequence 2, Application US/09264604
Patent No. 6346411
GENERAL INFORMATION:
APPLICANT: HOSTETTER, MARGARET K.
APPLICANT: GALE, CHERYL A.
APPLICANT: BENDEL, CATHERINE M.
APPLICANT: TAO, NIAN-JUN
APPLICANT: KENDRICK, KATHLEEN
TITLE OF INVENTION: CANDIDA ALBICANS GENE, INTEGRIN-LIKE
TITLE OF INVENTION: PROTEIN, ANTIBODIES, AND METHODS OF USE
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESSES:
ADDRESSEE: MUEITING, RAASCH, GEBHARDT & SCHWAPPACH, P.A.
STREET: 119 NORTH FOURTH STREET, SUITE 203
CITY: MINNEAPOLIS
STATE: MINNESOTA
COUNTRY: USA
ZIP: 55401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/264,604
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/642,846
FILING DATE: 03-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: MUEITING, ANN M.
REGISTRATION NUMBER: 33,977
REFERENCE/DOCKET NUMBER: 110.00280101
TELECOMMUNICATION INFORMATION:
TELEPHONE: 612-305-1217


```

? TELEPHONE: (619) 535-9001
? TELEFAX: (619) 535-8949
? INFORMATION FOR SEQ ID NO: 4:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 311 amino acids
? TYPE: amino acid
? TOPOLOGY: linear
?
US-08-460-309-4

```

Query Match	6.0%	Score 186.5;	DB 2,	Length 311;
Best Local Similarity	19.9%	Pred. No. 0.013;		
Matches 142;	Conservative 102;	Mismatches 256;	Indels 215;	Gaps 32

OY 28 NNOIJKKEE-----NNYGIEKRSBDRRESTVNDENAFILCKTPODNRHLLEE 78
 | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 1694 NEKIKLNETLIGTRDEAFEFNELEGLOK-----ETDOMIKELRRKNLETOKELIED 17433
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 OY 79 EVKSESSVESSNSSIDPAOQPSHTTN--REESVOTSDNVEDS-HVSDFANSKIKRESN- 133
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 1744 ELVAFAALKKVKKLLFCGESGNEEMEKDLREXLADYKKVDADWDLLEATDKIREANR 18030
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 OY 134 -----TSGSKE--ENTIEOPNKY-----KEDSTTS 156
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 Db 1804 LEAVNQNMATALEKKEAVEBSGRQJIENTLKESNDILDEANRLADEINSIIDYEVDIOPTK 18633
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 OY 157 QPSSGYTNIDEKISN-----ODELLNPINEYEKAKAPLTSTA-----OPSIRRYTVQOL 206
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 Db 1864 LPPHSELBNKDIDLGEIEKIDRLKLARKVSQAESHAAQLNDSASVALDIGLIDEKANINISFNAT 19233
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 OY 207 AAEDG-SNVANHLLKYVDOSTIBESIJDSEGIYAKHADENIITYVTIEVDDKYSAGDTMYDV 265
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 Db 1924 AAFAYAYINIKDYI-----DEAEKAK--EKKDLAHEAT-----KLATGR---C 1962Z
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 OY 266 IDKTVESDLDTSFTI-----PKIKDNSEGIATGT-----YFNKNKOIYTFPTYDV 313
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 Db 1963 LLKEDAGCCLOKSFRILINEAKLANLYKENEDHLNGLKRIENALARNODLLRTINDTLG 2022Z
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 OY 314 KYENI--KAHLKLTSYIDSKVPNNNTK-----LVEYKTALSVYNKITT 356
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 Db 2023 KLSAIPMDTFAKLOAVYDKARQANDPAKVDLAQITELHONLDGLTKKNYKRLADSVAKRNA 2082Z
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 OY 357 VEYOHPENNR-TANLOSMTFINID-----TKHNTYQGTYIYNPLRTS 396
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 Db 2083 V-VADPSPKNKITADADAVTNLBOEADRDLIDKPIKELEDNUKKNISISEIKELINOAKO 21414
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 OY 397 AKETNVNISGND-----EGS-----TIIDSTIIKY-----YK 425
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 2142 ANSIKIVASSOGDCIFRTYKPEIKKGSYNNIVWVKTAVADNLLFYLGSAKFIDFLATEMR 2201
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 OY 426 VGDMONTLPDSKRITDYSEYEDVIND-----YAOLGNNDNYNINFGNIDSPYLIKAYS 478
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 Db 2202 KGKVSFLMWDVGSGVREYEDLTIIDSZYWRIVASTRGNGTISVR--ALDOPKASIYPS 2259S
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 OY 479 -KYDPNKDDYTTOIQWTMOGTINEYTGEBERTASYDNITIAFSSTSOGGDGPPEPKYKI 537
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 Db 2260 THHSITSPGGTYIILDVANAMLPFGGLTGKRAKDAVARVITTFGCMEKETYPFDMK-----I 2314Z
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 OY 538 GDYVMEVDYDKDGIQNTNDNEKPLSNVLVTLLTYVD--GTISKSVRT--DEDGKYOFDG 569
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 Db 2315 G--LMNFERKEG-----DCKGCTVSPQVERSESSTIOFDG 2346
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

```

      NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESS: Campbell and Flores
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/125,077
FILING DATE: 22-SEP-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US PCT/US 94/10730
FILING DATE: 21-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/472,319
FILING DATE: 30-JAN-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/919,951
FILING DATE: 27-JUL-1992
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-LA 9721
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 3111 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-125-077-4

Query Match          6.0%; Score 186.5; DB 2; Length 3111;
Best Local Similarity 19.9%; Pred. No. 0.013;
Matches 142; Conservative 102; Mismatches 256; Indels 215; Gaps 32;

QY   28 NNOIIRKEET-----NNYDGIEKRSDEBTSESTVNDENAEFLQTPDNTHLTEE 78
DB   1694 NEKAIKLNETLTDRDAEFARNLEGJQ-----EIDQIKELRRKNLETQKEIMD 1743
QY   79 EYKESSSVESGSSSIDTYAOQPSHTTIN--REESVOTSDNVDS-HVSDFANSKRRESN- 133
DB   1744 ELVAEALIKKKYKLFGEISRGENSEMEKDLREKLADYKNKYVDAMDLLREAYDKIREANR 1803
QY   134 -----TESGKE--ENTIEGPNNV-----KEDSTS 156
DB   1804 LEAVANNKNMTALEKKKEAVESGRKOIENTLKGNIDILEANPLADEINSIIDYVEDIQK 1863
QY   157 QPSGYTNIDEKISN----ODELLNPINEYENKARPLSTSA-----QPSIKRVTVNOL 206
DB   1864 LPPMSEELNDKIDDLSQELTKDKRLKLEKYSQAASHAAQLNDSAVLDGLIDEAKNISFNAT 1923
QY   207 AAEGG-SVYNHILIKYTDOSITGEYGDDSEGCIVIAHA DENLIYDTVEVDKYSGDGMTVVD 265
DB   1924 AAFKAYSNIKDVI-----DEAEKVAK--EAKDLAHBAT-----KIATGPR---G 1962
QY   266 IOKNTVPSPDLTSEFY-----EKIDNGSEIIATGT---YNNKKQIITYFTFDYVD 313
DB   1963 LRKEPAKGCLQSFRILNEAKRLANDVKYKENEDHNLGLKTRIEMADARRGGDLLRTINDTLG 2022
QY   314 KYENI--KAHLKTSYIDSKVPPNNNTK-----LDVEKTTALSSVNKTIT 356
DB   2023 KLSAIPNDTAALKQAVKDKARAQNDTAKDVLAQITTELHONLDGLKKNYKRLADSVAKTNA 2082

```


GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 30, 2002, 10:28:55 ; Search time 52.53 Seconds

(Without alignments)
1084.732 Million cell updates/sec

Title: US-09-147-405-11

Perfect score: 3087
Sequence: 1 HHHHHHSSDDEKNDVYINN.....SKSVFDEGKQFDGQVVD 593

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : PIR_71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3015	97.7	1092	2 T30214	fibrinogen-binding
2	1146.5	37.1	1166	2 T32680	fibrinogen-binding
3	1139.5	36.9	1141	2 E89824	hypothetical prote
4	576	18.7	1315	2 T28679	fibrinogen-binding
5	559	18.1	1385	2 D89824	hypothetical prote
6	490	13.9	953	2 C89824	hypothetical prote
7	452	14.6	933	2 S41539	fibrinogen-binding
8	405.5	13.1	877	2 D89852	fibrinogen-binding
9	376	12.2	940	2 F90070	Clumping factor B
10	361	11.2	961	2 S19702	fibrinogen-binding
11	345.5	10.6	1038	2 H90053	hypothetical prote
12	326	9.8	1018	2 A32192	fibrinogen-binding
13	304	7.7	1039	2 T30856	protein F2 - strep
14	225	7.3	1463	2 T30229	AAS surface protei
15	222	7.2	1171	2 T18429	hypothetical prote
16	221.5	7.0	1093	2 AG3085	hypothetical prote
17	216.5	6.9	1125	2 E90598	membrane nuclease
18	214	6.9	1033	2 T18501	hypothetical prote
19	213.5	6.9	1033	2 T18275	1-phosphatidylinos
20	213	6.9	821	2 G71611	hypothetical prote
21	212.5	6.9	1125	2 E90598	hypothetical prote
22	211.5	6.9	1125	2 E90598	membrane nuclease
23	209	6.8	1127	2 T18427	hypothetical prote
24	208.5	6.8	1127	2 T18440	hypothetical prote
25	208	6.7	1072	2 A86827	hypothetical prote
26	207.5	6.7	1127	2 T28317	ORF MSV156 hypot
27	207	6.7	769	2 T28676	rhoptery protein -
28	207	6.7	769	2 F89870	serine proteinase
29	207	6.7	1939	2 T18372	repeat organellar

30	206	6.7	1302	1 J06009	surface-located me
31	205.5	6.7	1365	2 T30822	Impl protein - Myc
32	205.5	6.7	2510	2 T28160	hypothetical prote
33	205	6.6	2481	2 D90011	FmB protein (limp
34	205	6.6	3844	2 T18402	asparagine/asparta
35	203	6.6	537	2 A23770	asparagine-rich pr
36	202.5	6.6	1997	2 F71607	DNA helicase II BR
37	202.5	6.6	2269	2 T28677	rhoptery protein -
38	202	6.5	1308	2 E71622	probable membrane
39	201.5	6.5	4688	2 F82885	hypothetical prote
40	201	6.5	665	2 B71609	hypothetical prote
41	200.5	6.5	3147	2 T18674	hypothetical prote
42	200	6.5	2150	2 F71629	sensory transducti
43	199.5	6.5	1139	1 E64234	cytarabine-access
44	198.5	6.4	807	2 B71605	hypothetical prote
45	198.5	6.4	1805	2 T02712	similar to late em

ALIGNMENTS

RESULT 1
T30214
fibrinogen-binding protein - Staphylococcus epidermidis
C:Species: Staphylococcus epidermidis
C:Date: 22-Oct-1999 #sequence-revision 22-Oct-1999 #text-change 11-May-2000
C:Accession: T30214
R:Nilsson, M.; Frykberg, L.; Flock, J.I.; Pel, L.; Lindberg, M.; Guse, B.
Infect. Immun. 66, 2666-2673, 1998
A:Title: A fibrinogen-binding protein of Staphylococcus epidermidis.
A:Reference number: 220781; MUID:98261511
A:Accession: T30214
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1092 <NID>
A:Cross-references: EMBL:Y17116; NID:e1296734; PID:e1296735; PIDN:CAA76638.1

Query Match: 97.7%; Score 3015; DB 2; Length 1092;

Best Local Similarity 99.7%; Pred. No. 9,6e-122;
Matches 582; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY	8	SSDEKNDVYINNQSINTDNNQIKKEETNNVDGIEKSEDEETSTVNDENAEFLQK	67
DB	75	SSDEKNDVYINNQSINTDNNQIKKEETNNVDGIEKSEDEETSTVNDENAEFLQK	134
QY	68	TPQDNTHLTEEEYKSSVESSNSIDTAQOPSHTTINRESVQTSNDVESHVSDFANS	127
DB	135	TPQDNTHLTEEEYKSSVESSNSIDTAQOPSHTTINRESVQTSNDVESHVSDFANS	194
QY	128	KIESNTESGKEENTLEOPKPKVEDSTSPSGTYNDEKISNDELNLPINYEENKAR	187
DB	195	KIESNTESGKEENTLEOPKPKVEDSTSPSGTYNDEKISNDELNLPINYEENKAR	254
QY	188	PLSTSAOPSIKRVNTQALAEQSNVHLIKYTDOSITFGYDSEGVYKAHAENLIYD	247
DB	255	PLSTSAOPSIKRVNTQALAEQSNVHLIKYTDOSITFGYDSEGVYKAHAENLIYD	314
QY	248	VFEVDVKVSGDPTMTVDIDKNTVPSDLDSFTPIKIKNSGEIATGTYDNKKOITYT	307
DB	315	VFEVDVKVSGDPTMTVDIDKNTVPSDLDSFTPIKIKNSGEIATGTYDNKKOITYT	374
QY	308	FTDYVDKYNIAKHLKLTSTIDSKVPPNNNTKLDVEYKTAASSVNTTYEYRPNENRT	367
DB	375	FTDYVDKYNIAKHLKLTSTIDSKVPPNNNTKLDVEYKTAASSVNTTYEYRPNENRT	434
QY	368	ANQSMFTNIDTKNHVEQRTIYNPLRYSAKERNVNSGSGDGSITIDSTIIKYKYKG	427
DB	435	ANQSMFTNIDTKNHVEQRTIYNPLRYSAKERNVNSGSGDGSITIDSTIIKYKYKG	494
QY	428	DNQNLDSNRITDYSEEDVTNDYQALGNNNDVNIIFGNDISPYIIKYISKYDPKKDY	487
DB	495	DNQNLDSNRITDYSEEDVTNDYQALGNNNDVNIIFGNDISPYIIKYISKYDPKKDY	554

us-09-147-405-11.rpr

Query Match
Best Local Similarity 26.6%; Pred. No. 3,1e-12;
Matches 15; Conservative 10; Mismatches 234; Indels 86; Gaps 26;

A:Accession number: S41539; MID:94224142
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1,933 <MCD>
A:Cross-references: EMBL:Z18852; NID:g937525; PIDN:CAA79304.1; PID:g937526

14.6% Score 452; DB 2; Length 933;
8 SSDEKNDVINNNOSINTDDNQIIRKEETNNVDGIEKRSEDRTESTTNDENATFLQK 67
::: :::: :::: :::: -SDTKTSMTN--NGEITVAQ 91
Db 51 SNEKSNDSSVSAAPRTDNTV-----
68 TPQONTLLTEEVKSVESVSSSIDTAOOPSHTTINRESVOTSNVEDSHVDSANS 127
92 NP-----AOGETOSSSTNAATTETPYTGATTTTNQAOTPATPQS-SSTNAELVN- 143
Qy 128 KIKESNTSGKEENTIEDQPNKVKEDSTSPSGYTNIDEKISNODELIN-LPIYEENKA 186
::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::
Db 144 -QTSNETTFDNTNV-----SSVNSPQNSTNAEMVTQDTSTEATPSN---NES 189
Qy 187 RPLSTSQSPSIRKYTV-----QLAEQGSNNHLKTYDO--SITEGYDSEG 235
::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::
Db 190 APOSTDSNKDVNOAVNTSAPRRRAFSLAIVAADAPAAGTDITNLNTVVG1-DGGTT 248
236 IKADAENLIYDFEVDKKSGDTMTVIDDKNTVPSDLTDSFTIPKIXNSEGIATG 295
Date: 249 VYPHAGACVKNLNGFSPVNSAVKGDTEKITVPKELNLNGVSTFAKVPIIMG-QVLANG 307
Qy 296 TYDNKNKOITYTFDYDYKYNENIAHLKLISYIDSKVPPNNNTL-DVEYTAALS--VN 352
::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::
Db 308 VIDS-DGNVIYTFPDYVNTKDVRATLMTPAYID---PEWVKGTGNTLATYGISTAN 362
353 KITVEYORPBNRNTANLSMFNTIDFKNHVEQTIYINPL--RSARETNVINSIGNDE 410
::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::
Db 363 KTVLVDIRKKGKFKYLSTKGTIIDDKTNNTRYROTTYINPGDGVIAVLGNLKPNTDS 422
411 GSTIIDDSTIIKKVYKVGDNQLPDSNRILYDYSEYEDVTNDDYAOLGNNDVNIINF---- 466
Db 423 NALIDQNTSIKIVKYVDNAADLSESYFV-PNENEFEDVYSNVITFPNNQKVEENTPDD 481
467 NIDSYIIKIVSKXDPN-KDQYTITQOTVMQTINETGE--FRTAGYDNTIAFSTSG 523
Qy 482 QTTTYIYVNVGHIDPNSKGD-----LALRSLTYGINSNTIMSMWDNEVAFNNNSG 534
Db 534 OGCG-DLP--PEKTYKIGDY--VMEVDVKD-GIONTNDN 556
::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::
Db 535 SGDGIDKPVVPEQPDEPGEIERPEDSDSDPDSGSGS 573

RESULT 8
Db98852
fibrinogen-binding protein A, clumping factor [imported] - Staphylococcus aureus (str)
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: D89852
R:Kuruda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; G
ma, A.; Mizutani-Ut, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, H
C.; Shiba, T.; Hatori, M.; Ogawara, N.; Hayashi, H.; Hiramatsu, K.
L:Reference number: AB9758; MUID:21311952; PMID:11418146
A:Accession: D89852
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1,989 <KUR>
A:Cross-references: GB:BA000018; PID:g13700678; PIDN:BAB41975.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: cflA

Query Match
Best Local Similarity 26.6%; Pred. No. 3,1e-12;
Matches 15; Conservative 10; Mismatches 234; Indels 86; Gaps 26;

A:Accession number: S41539; MID:94224142
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1,933 <MCD>
A:Cross-references: EMBL:Z18852; NID:g397525; PIDN:CAA79304.1; PID:g397526

14.6% Score 452; DB 2; Length 933;
8 SSDEKNDVINNNOSINTDDNQIIRKEETNNVDGIEKRSEDRTESTTNDENATFLQK 67
::: :::: :::: :::: -SDTKTSMTN--NGEITVAQ 91
Db 51 SNEKSNDSSVSAAPTDDTN-----
68 TPQONTLLTEEVKSVESSESSSIDTAOOPSHTTINRESVOTSNVEDSHVDSANS 127
92 NP-----AOGETOSSSTNAATTETPYTGATTTTNOATPATPLOS-STNMDELVN- 143
Db 92 NP-----
128 KIKESNTSGKEENTIEDPKVKEDSTSPSGYTNIDEKISODELIN-LPIYEENKA 186
::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::
Db 144 -QSNETTFENDTNV-----SSVNSPQNSTNAEMVTQDSTEATPSN---NES 189
187 RPLSTSQSPSIRKYTV-----QLAEQGSNVHLKVTDO--SITEGYDSEG 235
::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::
Db 190 APOSTDSNKDVNOAVNTSAPRRRAFSLAIVAADAPAAGTDITNLNTVVG-I-DGGTT 248
236 IKADAENLIYDFVEVDKKSGDTMTVIDDKNTVPSDLTDSFTIPKIXDNSEGIATG 295
Date: 249 VYPHQAQGVKLNYGESPVNSAVKGDTEFIIVPKELNLNGVSTFAKVPIIMG-QVLANG 307
296 TYDNKNKOITYTFDYDYKYENIKAHKLISYIDSKVPANNNTL-DVEYTAALS--VN 352
::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::
Db 308 VIDS-DGNVIYTFPDYVNTKDVRATLMTPAYID---PEWVKGTGNTLATYGSTTAN 362
353 KITVEYORPMDNRANTANOSMFNTIDTFKNHVEQTYIYNPL-RYSARETNVINSNGDE 410
::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::
Db 363 KTVLVDIRKKGFKYLMSLTGKITDIQDKTNNTYROTIIYNPGDVAIAVLGNLKPNTDS 422
411 GSTIIDSTIIKYYKVGDNOLPDSNRILYDYSEYEDVTNDDYAQLGNNDVYNINF---- 466
Db 423 NALIDQNTSIKVVYVDNAADLSSESYFV-PNENFEDVYNSVNITFPNNQKVEENTPDD 481
467 NIDSYIIKIVSKXDPN-KDQYTIQQTVMQOTINETGE-FRTASYDNTIAFSTSG 523
482 QTTTPIYVNVGHIDPNKGD-----LALRSLTLYGNSNLTIMSMWDNEVAFNNNSG 534
534 OGQG-DLP--PEKTYKIGDY--VWEDVDKD-GIONTIND 556
::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::
Db 535 SGDGIDKPVVPEQPDEPGEIERPEDSDSDPDSGSGS 573

RESULT 8
fibrinogen-binding protein A, clumping factor [imported] - Staphylococcus aureus (str
Db9852
C:Species: Staphylococcus aureus
C:Update: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: D89852
R:Kuruda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; G
ma, A.; Mizutani-Ut, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, H
C.; Shiba, T.; Hatori, M.; Ogawara, N.; Hayashi, H.; Hiramatsu, K.
L:Reference number: AB9758; MUID:21311952; PMID:11418146
A:Accession: D89852
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1,989 <KUR>
A:Cross-references: GB:BA000018; PID:g13700678; PIDN:BAB41975.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: cflA

Query Match 14.6%; Score 452; DB 2; Length 989;

Best Local Similarity 25.9%; Pred. No. 3.4e-12;
Matches 152; Conservative 109; Mismatches 224; Indels 102; Gaps 27;

```

OY 8 SDEKNDVNNNOISINTDNNQIIKKEETNNVDGIEKSEDRTESTTVDNEAFLOK 67
   | : : : | : : | : : | : : | : : | : : | : : | : : | : : |
DB 51 SNESNSNDSSVSAAPKIDDTNV-----SDTKTSNTN---NGETVSAQ 91
OY 68 TPODTHLLEEEKSSSVSSNSIDTFAQPSHTTINREESVQSDNVESHVDFPANS 127
   | : : : | : : | : : | : : | : : | : : | : : | : : | : : |
DB 92 NP-----AQETQSSSNATTEEPVPGATTTTQANPATQOS-SNNAAELVN- 143
OY 128 KKESENTEGKEENTIEOPNKYKEDSTTSQPSGYTNIDEKISNOBELN-LPINEYENKA 186
   | : : : | : : | : : | : : | : : | : : | : : | : : | : : |
DB 144 --QTSNETTSNDNTNV-----SSVNSPQNSTNAENVSTQDTSTREATESN---NES 189
OY 187 RPLSTTSA-----QPSIKRVTVNOLA-----EAGSNVNLHKVTDQSITEGYDD 231
   | : : : | : : | : : | : : | : : | : : | : : | : : | : : |
DB 190 APONTDASKDVVSAVNPSTPRMRAFSLAAYAAPAAAGTDTN--QLTDVKVT---ID 244
OY 232 SEGVKRAHDAENLIYDVTFEVDVKVSGDTMTVDIDKNTVPSDLTDSFTIPKIDNGSEI 291
   | : : : | : : | : : | : : | : : | : : | : : | : : | : : |
DB 245 SCTTYVPHQAGYVKNLNGFSPNSAVKAGDTFKITVPEKELNGLVSTAKVPPIMAGD-QV 303
OY 292 IATGTYDNKKNQIITTFDYVYKYEIKAHKLSTIDSKVPNNNTKL-DVEYKAL-- 348
   | : : : | : : | : : | : : | : : | : : | : : | : : | : : |
DB 304 LANGVIDS-DGNVYITFDYVDNKEENVNTANITMPAYID---PENVTKGNVTLTGTIGT 358
OY 349 SSVNTIYEVORPENPNTANLOSMFTNIDTKNHTVEOTIYIPLKYSAKETV----- 402
   | : : : | : : | : : | : : | : : | : : | : : | : : | : : |
DB 359 NTKASTVLIDYEKGFHLSIKGFIQDIDKNTNNYRQIITVNP-----SGDNVVPALF 413
OY 403 -NISNGDEGSTIIDSTIIKYKVGDNQNLDPDSNRIYDYSEEDVTDNDYQAQGNNDV 461
   | : : : | : : | : : | : : | : : | : : | : : | : : | : : |
DB 414 GNLIPRTKSNALIDAKNFTIKYRV-DNANDLSESYVNPDEFELTVNQVRISFPANQY 472
OY 462 NINF-----GNISPIYIKIISKYDPKDKDYTIQCTVYMTTINETGEF--RTASYDMT 515
   | : : : | : : | : : | : : | : : | : : | : : | : : | : : |
DB 473 KVEFTDDQITTPYIVVNGHIDPASTG-----DLALRSTFYGYDSNFTWMSMDNE 526
OY 516 IAFSTSSGOGG-DLP--PEKTYKIDY--VVEDVDKD-GIONTNDN 556
   | : : : | : : | : : | : : | : : | : : | : : | : : | : : |
DB 527 VAFNNGSSGSDGIDKRVPEQDPPEIPIPEDSDSDSGSDSGSDS 573

RESULT 9
umping factor: B [imported] - Staphylococcus aureus (strain N315)
Species: Staphylococcus aureus
C:Accession: F90070
C:Date: 10-May-2001 #sequence, revision 10-May-2001 #text, change 22-Oct-2001
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogura, A.; Mizutani-Uji, Y.; Kobayashi, N.; Savano, T.; Inoue, R.; Kato, C.; Sekimizu, K.; C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.; Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PMID:11418145
A:Accession: F90070
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-877 <KUR>
A:Cross-references: GB:BA000018; PID:g13702588; PIDN:BAH43728.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: clfB

Query Match 13.1%; Score 405.5; DB 2; Length 877;
Best Local Similarity 24.4%; Pred. No. 2.8e-10;
Matches 144; Conservative 82; Mismatches 210; Indels 153; Gaps 23;

```

```

OY 4 HHHSSDEEKNV---NNOSINTDNNQIIKKEETNNVDGIEKSEDRTESTTVDNE 60
   | : : : | : : | : : | : : | : : | : : | : : | : : | : : |
DB 39 NHOQAQSEDSNDTTOSSKNNAASADSEKNMT-----ETPOLNTANTDSISANTNAND 94
OY 61 EATFLOKTPDQDTHLLEEEKSSSVSSNSIDTFAQPSHTTINREESVQSDNVESH 120
   | : : : | : : | : : | : : | : : | : : | : : | : : | : : |
DB 95 SPTKPMSTQTSMTTTE-----PASTNETPQPAIKNQ---ATAAKMODT 137
OY 121 VDFANSKIKESNTESGKEENTIEOPNKYKEDSTTSQPSGYTNIDEKISNOBELNLPIN 180
   | : : : | : : | : : | : : | : : | : : | : : | : : | : : |
DB 138 VPOEANSQVQDNKTT---NDANSIATNSELKNSQT-----LDLP-- 172
OY 181 EYENKARPLSTSAO---PSIKRVTVNOLAEO-----GSNVNLHKVTDQSITE 227
   | : : : | : : | : : | : : | : : | : : | : : | : : | : : |
DB 173 ---QSPPTISNAQTSKPSVTRAVBSLAAPVYVNAADAKGTNVN-----DKVTAS 222
OY 228 GYDSEGVKRAHDAENLIYDVTFEVDVKVSGDTMTV-----DID---KNTPVS 273
   | : : : | : : | : : | : : | : : | : : | : : | : : | : : |
DB 223 NFKLEKTFDPNQSNTFMAANFTVTDKVGSGDYFAKLPSLTLGVDYDSNNTMP- 281
OY 274 DLTDFTPIKIDNSELATATGTDNKNQIITTFDYVDKYEIKAHKLSTIDSKVY 333
   | : : : | : : | : : | : : | : : | : : | : : | : : | : : |
DB 282 -----IDIKSTNDGVAKATYDILTKTYTFVFTDVNNMENINGQFSLPLETDRAA 334
OY 334 PNNNTKLDVEYKTAALSVNKTTFVEYORP-----NENRTANLOSMFTNIDTKN--HTVEQT 387
   | : : : | : : | : : | : : | : : | : : | : : | : : | : : |
DB 335 PKSGT-YDANINADMFNKKITVNTSSPIAGIDKNGANISSQIIGVYTAAGQNTYKQT 393
OY 388 IYINPLRYSAKETNNVINSNGDEGSTIIDS-----TIKRVKVGDNQLPDS--- 435
   | : : : | : : | : : | : : | : : | : : | : : | : : | : : |
DB 394 VFNPNKQVILGNTWVYIKGYOK-----IESSGKVASATPKLRIFEVNDTSKLSDSYAD 449
OY 436 -----NRIYSEYEDVTDNDYQAQGNNDVNINFGNDSPIIHKVSKVD 481
   | : : : | : : | : : | : : | : : | : : | : : | : : | : : |
DB 450 PNDNLKEVTDOFKNRY---YE-----HPNVAIKFGDITTYVVLVLEGHTD 494
OY 482 PKNDYTT--IQTVMQTTINEYTGEPRTASYDNTIAFSTSSGOGGD 528
   | : : : | : : | : : | : : | : : | : : | : : | : : | : : |
DB 495 NGCKNKLTQVIGQNDVPV-----NRDYSIFGNNNENVRYGGSGSADGD 538

```

```

RESULT 10
S19702
fibronectin-binding protein B - Staphylococcus aureus
C:Species: Staphylococcus aureus
C:Date: 13-Jan-1995 #sequence, revision 13-Jan-1995 #text, change 15-Oct-1999
R:Joensson, K.; Signaes, C.; Mueller, H.P.; Lindberg, M.; Eur. J. Biochem. 202, 1041-1048, 1991
A:Title: Two different genes encode fibronectin binding proteins in Staphylococcus au
A:Reference number: S19702; MUID:92111475
A:Accession: S19702
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-940 <JOE>
A:Cross-references: EMBL:62992; NID:g49040; PIDN:CAA44726.1; PID:g581562
C:Keywords: fibronectin binding

```

```

Query Match 12.2%; Score 376; DB 2; Length 940;
Best Local Similarity 23.4%; Pred. No. 5.6e-09;
Matches 138; Conservative 110; Mismatches 217; Indels 124; Gaps 22;

OY 42 GIKRSEDREESTTNDENATFLQKTPDQDTHLLEEEKSSSVSSNSIDTFAQ----- 97
   | : : : | : : | : : | : : | : : | : : | : : | : : | : : |
DB 30 GQKEAASEQNNNTVEES-----GSSATSKASSETQTTNNVNTIDETQSYSA 78
OY 98 ---OPSHTP-IRRESVOT--SDNVESHVDFPANSKIKESNTESGKEENTIEOPNKVY 150
   | : : : | : : | : : | : : | : : | : : | : : | : : | : : |
DB 79 TSTEOPSQSTQVTTTEAPKIVQAPEVETSNV-DLPSEKVAADKETTGYDD--IAQTSNVS 135
OY 151 EDSTTSOPSGYTNIDEKISNOBELNLPINEYENKARPLSTSAQPSIKRVTVNOLAEO 210
   | : : : | : : | : : | : : | : : | : : | : : | : : | : : |

```

Db 136 E-----IKPRMRSQDVTVAEK 153
QY 211 -----GSNNHLIKYTDOSITBEGYDSEVYKAHAENLIYDTEFVDKVKSGDT 261
Db 154 EVEETKATCTDVTKNVEEVEEGSEIVGHKODTAVVPHNNEFVTLKKMKFEGIKAGDY 213
QY 262 MIVDIDKNTVPDLSFTIPKIKDNGSEIATGYDNKKKCLTYPTDQVDDYENIKAH 391
Db 214 FDTTLDONVETHGISILRKVPEIKSTDQVMAIGEIIIGENK-VRYTFKEVQEKDULAE 272
QY 322 LKLTSLYIDSKVPPNNTKLDVEKXTALSYNKTITVEY---QRPNNRNTANLQSMFTNID 378
Db 273 LSLNLFIDPTTVYQKGNQ-NVEVKLEGETVSKIFENIQYIGVADNMGVLAN-----GRID 326
QY 379 TKNHT---VEQTIYINPLRSKAKETVNVISGNDGEGTIIDSTIIKRVKGVGNOLPDS 435
Db 327 TLNRKVDGKFSHFAVYKPRNNQSL--SSVTVYGOVTKGKRKPGVNPYKVVYHHISDLSAES 384
QY 436 --NRIYDSEYEDVTNDYDQAOLGNNDNINFGNIDSP--YIIKRVISKYDPNKKDDYTTIQ 491
Db 385 VYAKLDDVSKFEEDVTNMSLDPTNGSYSLNFNNLDQSKRYVYIKYEGYDYSNAN-----439
QY 492 QVVTMOITINETYGEFRTA--SYDNTIAFSTSSGQGG-----DLPEKTYKIG 538
Db 440 --LEFQTHLFGYNYYYTSLNLTWKNGVAFYSNNAQGGDKLKEPIIEHSTPIELEPKSE 497
QY 539 DYVMEVDKQCIQNTDNEKPLSNVLVTLT-----YPDGISKVRIDED 582
Db 498 PVEKEHLEGTIEESNDS-KPIDFEYHTAVEGAEGHAEGT---IETED 542

RESULT 11

G90053
Hypothetical protein fnbB [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #extl_change 22-Oct-2001
C:Accession: G90053
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc
ma, A.; Mizutani-U, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiratsuyu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: G90053
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1961 <KUR>
A:Cross-references: GB:BA000018; PID:g13702452; PIDN:BA843593.1; GSPDB:GN00149
Experimental source: strain N315
C:Genetics:
A:Gene: fnbB

Query Match 11.28; Score 345.5; DB 2; Length 961;
Best Local Similarity 22.2%; Pred. No. 1.2e-07;
Matches 125; Conservative 132; Mismatches 219; Indels 87; Gaps 24;

QY 81 KESSVSSNSIDTAQPSHTTINRESVQTSQVDSVSDPANSKIKESNTESKEE 140
Db 33 KEAASQNNNTYVESGS-SATESKASETQTTNNV-----NTIDETOSYSA---78
QY 141 NTEOPKRVKEDSTTSQPSGY--TNIDEKISNDELINLPINEYENK-----ARPLSTP 192
Db 79 TSTPEKPSQVTEEAFTVQAPKVFTEKMSQEDLPSEKVADETTGTQVDAQPSNVS 138
QY 193 SAQPSIR-----VYVNOJLAAE--GGSNNHLIKYTDOSITBEGYDSEGVYKAHAEN 243
Db 139 EIKPRMRSADVAVSEKEVAEKAATGDTNKNVEYTESSL-IGHNKKDSNIVAPHNAQR 197
QY 244 LIYDVEEDVDKYSQGTMTVVDIDKNTVPBDLIDSTFIPIKIDNS-GEIATGTVDYDKK 302
Db 198 VTLKMKFEGIKAGDYFDTLSDNVEHGISILRKVPEIKSTDQVMAIGEVINE-R 256

QY 303 QITVFEDVDYKXENIKAKHLKLSYIDSKVPPNNTKLDVEKXTALSYNKTITVEYQRP 362
Db 257 TIRYPTDITNNKDLTAEINLNLFTDPTTVYKQSGQ-KVEYTLGQKVKSEPDYLLDG 315
QY 363 NENR---TNLQSMFTNIDTKNH--TVEQTIYINPLRSKAKETVNVISGNDGEGTIID 416
Db 316 VKRMGVTVN-----GRIDLKEEKGFSHFAVYKPRNNQSL--TSVTVYGOVYSGVKQSA 368
QY 417 DSTIITVYKGVGNOLPDS--NRIYDSEYEDVTNDYDQAOLGNNDNINFGNIDS--PY 472
Db 369 NNPTVYKHHIGSDELAESYAKLDTSKFEEDYTERVNLSTYNGGTYLMLGDLNDSKD 428
QY 473 IIVISKYDPNKKDDYTTIQOTVMTQITINETYGEFRTA-----DNTIAFSTSS 522
Db 429 VIKYGEYDQNAKD-----LNFRTHLGSHKTYPIYPIYPIYPIYQVLTNNKVAFTSN 481
QY 523 GQGGQ-DLPEKTYKIGDYVMEVDKRD-----GIQNTDNEKPLS-NVLVTLTYP 570
Db 482 AKGDGKRPNDPIEKESEPIDLIKSEPPYERKEHLEGTIEESNDS-KPIDFEYHTAVEGA 540
QY 571 DGTSKSVRIDEDGKYQPDQVQVD 593
Db 541 EGHAEGLIETEE-----DSIHVD 558

RESULT 12

H90053
Hypothetical protein fnb [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #extl_change 22-Oct-2001
C:Accession: H90053
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; C
ma, A.; Mizutani-U, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiratsuyu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: H90053
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1038 <KUR>
A:Cross-references: GB:BA000018; PID:g13702453; PIDN:BA843594.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: fnb

Query Match 10.68; Score 326; DB 2; Length 1038;
Best Local Similarity 23.1%; Pred. No. 8.8e-07;
Matches 139; Conservative 102; Mismatches 215; Indels 146; Gaps 30;

QY 42 GIKRSEDRTESTNTNDENATFLQKTPQDN-----THLTPEEYKESSEVSSNSIDTAQ 97
Db 30 GQDKKAASEQKTYVEEN-----GNSATDNKTSETQTTNNHLEETOSYNATV--TE 82
QY 98 QPSHTT-INRESVQTSQVDSVSDPANSKIKESNTESKEENTIEQPNKYKEDSTTS 156
Db 83 QPSNATQVTEEAQKAVQ-----APQTAQPAVVEYKKE--EK-QVKE--TT 125
QY 157 QPSGYTNIDEKISNDELINLPINEYENK-----ARPLSTSAQPSIKRYVNOJLAA 208
Db 126 QP-----QNSGNQNRQVDTLPKVVQNOQGTQVEVAQPRTASESKPRVTR-SADVAEA 178
QY 209 EGSNNHLIKYVDOSTBEGYDSEGVYKA-----HAENLIYVTEFVDKVKSGDT 260
Db 179 KEASDVSE-VKGD--VTSKTVYESGSEIAPQGNKYEPHAGQRVNLKYLKFPADGLKRGD 235
QY 261 TMTVVDIDKNTVPBDLIDSTFIPIKIDNS-GEIATGTVDYDKK 315
Db 236 YPFTLSNNNTYGVSTARKVPEIKNSVYMAIGEILGN-----NIRYTTNTELEHK 288
QY 316 ENIKAKHLKLSYIDSKVPPNN-----NTKLDVEKXTALSYNKTITVEYQRPNNRTANL 370

Db 289 VEVANLEINLEIDPKTVQNGEOKITSLNGE-----ETEKTIPVYNPGVNSNTNV 342
QY 371 QSMETNIDTNRHVEQITINPLAYSARENVNINSGNDESTIITDSTIIKYK-VGDN 429
Db 343 NGSIETPNKSNKFTTHAIYIKPMNGNSMT-VSVTGLTLESSLNLAGOPVYKVEYLCK 401
QY 430 ONLPDSNRIT---DYSEEDVTNDYQAOLG--NNNDVINFGNIDSPYIITKISKYDPN 483
Db 402 DELPQS--VVAISDITKFKDYTKEMNGKLSVQNGSYSLMDKIDTYVHYHGEYLQ 459
QY 484 KDPYTTIOQTVMQITINEYTGFEPTASVDTNIAFSTSSGOGDLPPEKTYKIGDYWE 543
Db 460 SDQY-----NRTELYGY-----PERAYK-SYVYV- 483
QY 544 DVKDGIONTNDNEKPLSNVLTLYTPDGTSKS--VPTDEGKY-----QPDGV 590
Db 484 ---GGYRLTWDN-----GLVLYSNKADGNKNGOITODNDFEYKEDTAKTMSGOYDAK 534
QY 591 QY 592
Db 535 QY 536

RESULT 13
A32192
fibronectin-binding protein - Staphylococcus aureus
C:Species: Staphylococcus aureus
C:Date: 31-Jul-1989 #sequence_revision 31-Jul-1989 #text_change 16-Feb-1997
C:Accession: A32192
R:Signes, C.; Raucel, G.; Joensen, K.; Lindgren, P.E.; Anantharamiah, G.M.; Hoeck, M.
Proc. Natl. Acad. Sci. U.S.A. 86: 699-703, 1989
A:Title: Nucleotide sequence of the gene for a fibronectin-binding protein from Staphylococcus aureus
A:Reference number: A32192; MUID:8908998
A:Accession: A32192
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1018 <SIG>
A:Cross-references: GB:J04151
C:Keywords: fibronectin binding

Query Match 9.8%; Score 304; DB 2; Length 1018;
Best Local Similarity 21.2%; Pred. No. 7.5e-06;
Matches 12; Conservative 12; Mismatches 24; Indels 104; Gaps 26;
QY 42 GIERSEDEPTSTNVDENEATFLQKTPDN---THLLEEVKSSSVSSNSIDTAQ 97
Db 30 GQDEMAASEQKTTVEEN-----GNSATDNKTSOTATTANNHIEFTQSYNATV--TE 82
QY 98 QPSHTT-INREE---SVQTSNVEDSHVDFANSKIKESNTESGKEENTIQPNKVKEDS 153
Db 83 QPSNATQVTEEAARKAVQAPOTAPANIEYKVEYKE-----EAKPKVKETT 130
QY 154 TTSQPSG---TNTIDKISNDELINLPINEYENKARLSTSAQESIKRYT-----V 203
Db 131 OSQNSGGQROVDLPKPKATONVAETQV---EVAQPTASSESKFRTSRSDAFAEKA 186
QY 204 NQLAEOGSNNVHLIKYDOSITEGYDSEGIKAHAENLIYDTFVDDKVKSGDPTMT 263
Db 187 SNAKETSTGTVTSKYVETIGSI-BGHNNKV-EPHAGQRAVLAKKLFPENLHGQDFD 244
QY 264 VDIKNTYPSDLTDSFTLPKIKDMS---GEIATGYDNKKNKOITTFDDYVDKYEVI 318
Db 245 FTLSNNVTTHGVSTARKVPEIKNGSVNATGEVLEGG-----KIRYFTNDIDKVDY 297
QY 319 KAHKLSTYIDKSKVNNNTKLDVEYKTALSSVN-----KITVEYCPNENRTANLOSM 373
Db 298 TAELEINLFIDPKTYQTNQ-----QTTITSLNEQSKELDVKYKQSIGVYANLNGS 351
QY 374 FTNIDTKNHTVEQITINPLAYSARENVNINSGNDESTIITDSTIIKYK-VGDNOL 432
Db 352 IETFKANNRSHVAFIKP--NNGKTSVTVYGLTKMGSNQNGQPKVRIEFLNNDI 409

QY 433 PDSNRIT---DYSEEDVTNDYQAOLG--NNNDVINFGNIDSPYIITKISKYDPND- 485
Db 410 AKS--VVAISDITKFKDYTKEMNGKLSVQNGSYSLNIEHLDKIDTYVHYHGEYLQ 467
QY 486 DVYTTIOQTVMQITINEYTGFEPTASVDTNIAFSTSSGOGDLPPEKTYKIGDYWE 544
Db 468 VDFET--QWVGHPEQLKYKYDRGYTLTWDNGLVLYSNKANGN----- 508
QY 545 VDKRG--IONTNDNEKPLSNVLTLYTPDGTSKSVRDEDEKQYQPDGVQVD 593
Db 509 -EKNGPIION-NKPEYKEDTAKERLT--GOYDKMLVTTVEEYDSSLDID 555

RESULT 14
T30856
protein F2 - Streptococcus pyogenes
C:Species: Streptococcus pyogenes
C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 11-May-2000
C:Accession: T30856
R:Jaaffe, J.; Natanson-Yaron, S.; Caparon, M.G.; Hanski, E.
Mol. Microbiol. 21, 373-384, 1996
A:Title: Protein F2, a novel fibronectin-binding protein from Streptococcus pyogenes
A:Reference number: 220907; MUID:97011581
A:Accession: T30856
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1039 <AAF>
A:Cross-references: EMBL:U01980; NID:G1654115; PID:G1654116; PIDN:AA04522.1
C:Genetics:
A:Note: PRF2

Query Match 7.7%; Score 239; DB 2; Length 1039;
Best Local Similarity 22.6%; Pred. No. 0.0046;
Matches 93; Conservative 79; Mismatches 168; Indels 72; Gaps 18;
QY 224 SITEGYDSEGVIRAHND--AENLIYDTEFEVDKVS-----GDTMYDIDKNTVPSDL 275
Db 38 SVSKRYGEQERTSNADRYNNHAAVFKMSFELKOKDSSETINPDTEYVQLDRRLNPKGI 97
QY 276 TDSFTIPK-INDNSGELIATGYDNKKNKOITTFDYDYKYNIAHLKITSYIDSKVP 334
Db 98 SOD--IRKIIIDSENSPLAIKIDAKTHQLYTTNTVYAGIDKYQLSLELPLENKEVL 155
QY 335 NNNKRLDVEYRTALSSVKNKTTVEYQRPNE-----NRTANLOSMFTNIDKKNHVE 385
Db 156 ENFNISDKSTIGQETIYKGTAVNLYGNESKESNTYTNGLSNVSGSIESYNETGEFV 215
QY 386 QTIYINPLAYSARENVNINSG-----NGDEGSTIIDS--TIKYYKVGDNQNPDSNR 437
Db 216 WYVYVNPRTNIPYAVLMLMGFAKRTAOGENDNSLVSAOGLGYDIEVPHNYRLPTSYG 275
QY 438 IYDSEYEDVTNDYQAOL-----GNNDVININFG-NID-SPYIITKISKYDPNDVYTT 489
Db 276 V-DISRL-NLKKDLAKLPQSGTQGANRKLRIIDRENIQGAFAVYKVGKADQSKF--- 330
QY 490 IOQVYTMQITINEYTGFEPTASVDTNIAFSTSSGOGG-----DLPEKT 534
Db 331 ---LIVQSHLSFNNMGSYKILRPNSHVSEFTNEIALSPKSGSGSTSEFTKPSITVANL 386
QY 535 YKIGDYVEDYDKDI-----QNTNDNEKPL--SNVLTLYTPDGTS 574
Db 387 KRVAOLRFKKVSTQVNPVLEPAAFELRLSSNGNSOKLEASSNTOGETHFRDLTS 438

RESULT 15
T30290
AAS surface protein - Staphylococcus saprophyticus
C:Species: Staphylococcus saprophyticus
C:Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 02-Sep-2000
C:Accession: T30290
R:Hell, W.; Meyer, H.G.W.; Gattermann, S.G.
Mol. Microbiol. 29, 871-881, 1998

A:Title: Cloning of aas, a gene encoding a Staphylococcus saprophyticus surface protein
A:Reference number: Z20809; MUID:96389318
A:Accession: T30290
A:Status: Preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1463 <HEL>
A:Cross-references: EMBL:AJ000007; NID:e1295629; PID:e1295630; PIDN:CAA03852.1
C:Genetics:
A:Gene: aas

Query Match: 7.3%; Score 225; DB 2; Length 1463;
Best Local Similarity 20.9%; Pred. No. 0.029;
Matches 143; Conservative 105; Mismatches 263; Indels 172; Gaps 33;

```
OY      8 SDEKNDVNNNOI-----NTDNNQIIRKEETNNYIGIEKSEDRTESTT--- 55
DB      180 NTDENSDVANNQNEPVAQNDKAETSNEDVASSDVAKQDGTSDNANASDLTDQNESETLND 239
        56 -NYDENAEFLQKTPODNH-----LTE--EEVKSSSVSSN---SIDTAAQPSH 101
DB      240 NAVSSNEDVASSDVAKQDGTSDNANASDLTDQNESVAQNDKAETSNEDVASSDVAKQDDTH 299
OY      102 TTIN-----REESVQTSQDNVEDSHVDFANKIKESNTESG-----KEENTIEOPN- 147
DB      300 SPANASDVADQNESETQNDNMESSNEDVASSDVAKQDDTHSDANASDVAKQNESETQND 359
OY      148 -----KVKEDSTS--QPSGYTNIDEKISNODELLNLPINEYENKAPLSTT 192
DB      360 KAETSNEDVASSDVAKQDGTSDNANASDLTDQNESVAQNDKA-----ETSNEDVASS 411
OY      193 SAQPSIKRTVAVQGLAEGGSNNHLIKYT--DQSITEGYDS---EGYIK-AHDAENLI 245
DB      412 DKQDDTHSDANASDLTDQNESATQDDKAISKEDDVASNDKQDNKAVSNIIKEASTAENKV 471
OY      246 YDVT--EYVDKVKSGDPM--TVDIDKNTVPEDLTFDSFTPIKIDNSGEIATGTYDNKN 301
DB      472 QPRTSAKATPKLRATVATSANTAVATRSATVTEATTRALPKY-----S 515
OY      302 KOITTFDYDYKVENIKA--HLKLTYSIDSKVPPNNNTKLADVEYKTALSSVNKTIVE 358
DB      516 PKVNSSINNYIRK--NNFKAPNEYQDIANYLPQYNRYRGKPEGIYMHDTA--NDNSTITGE 572
OY      359 YQRPNEPNTANQSMFTNIDTNHVEQIYINPLRYSAKETNWNISGNGDEGSIIDDS 418
DB      573 INYMKNNYT---SAFVAHYVDGDRILETANTDYLAWGA-----GPOANDRF 615
        419 TIKYKYVQDNQNLDPDSNRIYDYSEVEDTNDVYQALGNNNDVNFNGNIDSPYIIKYIS 478
DB      616 IHELVTHTHDYDFARS--INNYADYA--ATNQYIGLVDP---SAEYDGVGTWTHQAVS 669
OY      479 KY-----DPN-----KDYTTIQQTVMTQTTINEYTGERTASYDNTAFST- 520
DB      670 NYLGSSDSDPHGYLAHNYSYDELYDLIEKYLITKG-----QAAAGTSSGSGTG 721
OY      521 ----SSGQGO--GDLRPEK--TYKIGDYWEDVDKGIQNTNNEKPLSNVLTLYPPDGT 573
DB      722 GTGGSTGSGNTGTPPSKSGYKV-----TENNGVGRINSK---NDGLTYTVYDQKG 770
OY      574 SKSVRTDED-----GKYOF 587
DB      771 KTTDRVNOQLKTKTSATLGKEQY 793
```

Search completed: July 30, 2002, 10:31:29
Job time: 154 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 30, 2002, 10:31:56 ; Search time 28.43 Seconds

(without alignments)
807.622 Million cell updates/sec

Title: US-09-147-405-11

Perfect score: 3087
Sequence: 1 HHHHHHPSDEKNDYINNN.....SKSVFDEDKYQFGVQVD 593

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database: SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES:

Result No.	Score	Query	Match Length	ID	Description
1	304	9.8	1018	1	FNBA_STAU
2	214	6.9	1093	1	PI4K_DICD1
3	203	6.6	537	1	ARP_PLAFA
4	199.5	6.5	1139	1	HMM1_MYGE
5	197	6.4	918	1	YMB_CAEEL
6	195	6.3	731	1	BAF1_YEAST
7	192.5	6.2	578	1	LIPA_MYCPU
8	189.5	6.1	1664	1	INT1_CANAL
9	188.5	6.1	1435	1	EBAL_PLARC
10	186.5	6.0	3110	1	LMA2_HUMAN
11	186	6.0	1381	1	YB7_YEAST
12	185.5	6.0	1165	1	YNF4_YEAST
13	184.5	6.0	1744	1	TANA_XENLA
14	184	6.0	817	1	YGA4_YEAST
15	184	6.0	1251	1	RRP2_PLAVB
16	184	6.0	1658	1	YMG7_YEAST
17	183.5	5.9	2009	1	SEC7_YEAST
18	183.5	5.9	2452	1	RPL1_PLAFD
19	182.5	5.9	853	1	YCG1_YEAST
20	182.5	5.9	1875	1	MLP1_YEAST
21	182	5.9	1419	1	ALAI_CANAL
22	182	5.9	1460	1	NI59_YEAST
23	181.5	5.9	1093	1	YKD5_CAEEL
24	180	5.8	1253	1	DSP6_HUMAN
25	178	5.8	960	1	YMX6_YEAST
26	177.5	5.7	1196	1	BXCN_CLOBO
27	176.5	5.7	1637	1	MRSP_STAU
28	176.5	5.7	1803	1	YUL3_YEAST
29	175.5	5.7	1183	1	CNA_STPAU
30	175	5.7	1147	1	CGA1_HELPY
31	174.5	5.6	682	1	NI5P_LACLA
32	174	5.6	1466	1	SPAZ_YEAST
33	174	5.6	2022	1	ANP1_ONCVO

34	173.5	5.6	967	1	Y5G0_CLOAB	P33747 clostridium
35	173.5	5.6	1260	1	ALSI_CANAL	P46590 candida alb
36	173	5.6	827	1	CSG_HALVO	P25062 halobacteri
37	172.5	5.6	688	1	LIP_STARP	P02510 staphylococ
38	170.5	5.5	970	1	PSU1_YEAST	P33550 saccharomyc
39	170.5	5.5	1435	1	LTEL_YEAST	P07866 saccharomyc
40	170.5	5.5	1636	1	BUD3_YEAST	P25558 saccharomyc
41	170	5.5	2334	1	WAPA_BACSU	P07833 bacillus su
42	169.5	5.5	571	1	TACY_STRPY	P21131 streptococc
43	169.5	5.5	1020	1	BCA_STRAG	P002197 streptococ
44	169.5	5.5	2869	1	RBP1_PLAVB	P00799 plasmodium
45	169	5.5	954	1	BRL1_YEAST	P47134 saccharomyc

ALIGNMENTS

RESULT	1	FNBA_STAU	STANDARD;	PRT; 1018 AA.
ID	FNBA_STAU			
AC	PI4738:			
DT	01-APR-1990 (Rel. 14, Created)			
DT	01-APR-1990 (Rel. 14, Last sequence update)			
DT	30-MAY-2000 (Rel. 39, Last annotation update)			
DE	Fibronectin-binding protein precursor (FNBP).			
GN	FNBA.			
OS	Staphylococcus aureus.			
OC	Bacteria; Firmicutes; Bacillus/Clostridium group;			
OC	Bacillus/Staphylococcus group; Staphylococcus.			
OX	NCBI_TaxID=1280;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=NCYC 8325-4;			
RX	MEDLINE=89098998; PubMed=2521391;			
RA	Signaes C., Raucet G., Joensson K., Lindgren P.-E.,			
RA	Anantharamiah G.M., Hoeek M., Lindberg M.;			
RT	"Nucleotide sequence of the gene for a fibronectin-binding protein			
RT	from Staphylococcus aureus: use of this peptide sequence in the			
RT	synthesis of biologically active peptides".			
RL	Proc. Natl. Acad. Sci. U.S.A. 86:699-703(1989).			
CC	-1- FUNCTION: THE ABILITY OF BACTERIA TO BIND FIBRONECTIN HAS BEEN			
CC	PROPOSED AS A VIRULENCE FACTOR ENABLING BACTERIA TO COLONIZE			
CC	WOUND TISSUES AND BLOOD CLOTS. BINDING OF PLASMA FIBRONECTIN TO			
CC	THE BACTERIAL SURFACE MIGHT BLOCK ADHESION RECEPTORS ON S.AUREUS,			
CC	THIS REPRESENTING AN IMPORTANT DEFENSE MECHANISM AGAINST TISSUE			
CC	INVASION.			
CC	-1- SUBCELLULAR LOCATION: Type I membrane protein. Cell wall.			
CC	-1- SIMILARITY: TO OTHER STREPTOCOCCAL AND STAPHYLOCOCCAL PROTEINS			
CC	IN THE REGION OF THE MEMBRANE ANCHOR.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation-			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@sib-sib.ch).			
CC	-----			
DR	EMBL; J04151; AAA26632.1; -			
DR	InterPro: IPR004237; Fn_bind.			
DR	InterPro: IPR001899; Gram_pos_anchor.			
DR	Pfam: PF00746; Gram_pos_anchor. 1.			
DR	Pfam: PF00746; Gram_pos_anchor. 1.			
DR	PROSITE; PS00343; GRAM_POS_ANCHORING. 1.			
KW	Signal; Repeat; Cell wall; Transmembrane.			
FT	SIGNAL	1	36	FIBRONECTIN-BINDING PROTEIN.
FT	CHAIN	37	993	EXTRACELLULAR (POTENTIAL).
FT	DOMAIN	994	1009	MEMBRANE ANCHOR.
FT	TRANSMEM	994	1009	CYTOPLASMIC (POTENTIAL).
FT	DOMAIN	1010	1018	B-1.
FT	REPEAT	545	574	B-2.
FT	REPEAT	575	604	4 X APPROXIMATE TANDEM REPEATS,
FT	DOMAIN	745	878	


```

FT REPEAT 745 782 FIBRONECTIN-BINDING DOMAIN.
FT REPEAT 783 820 D-1.
FT REPEAT 821 859 D-2.
FT REPEAT 860 878 D-3.
FT REPEAT 879 948 D-4 (INCOMPLETE).
FT REPEAT 879 948 5 X TANDEM REPEATS, PRO-RICH (WR).
FT REPEAT 879 892 WR1.
FT REPEAT 893 906 WR2.
FT REPEAT 907 920 WR3.
FT REPEAT 921 934 WR4.
FT REPEAT 935 948 WR5.
FT DOMAIN 982 987 CONSERVED IN GRAM-POSITIVE COCCI SURFACE
                          PROTEINS.
SQ SEQUENCE 1018 AA; 111780 MW; 58175E002081F1F CRC64;

Query Match          9.8%; Score 304; DB 1; Length 1018;
Best Local Similarity 21.2%; Pred. No. 4.7e-06;
Matches 125; Conservative 121; Mismatches 241; Indels 104; Gaps 26;

42 GLEKSEDETESTNDENEAFLQKTPDN---THLTEEYKSSSVSSSSSIDTQ 97
| : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : :
30 GDDKKAASEQKTYVEEN---GNSANDNKTSETOTATVNHIEQSYNATV--TE 82
| : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : :
98 OPSHTT-IRRE---SVOTSDNVEDSHVDFANSKIKESNTSGKEENTIEQPNKEDS 153
| : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : :
83 QPSNATQVTEAPRAKAVAPQTAOPANITVKEEYVKE-----EAKPOYKETT 130
| : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : :
154 TTSQPSG---YTNIDEKISNOELLNLPINEXENKARPLSTSAOPSIKRVT-----V 203
| : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : :
131 OSQDNGSGDQROVLPKPKATQNOVAETQV---EVAQPTASESKRPVTRSAADVAEAKA 186
| : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : :
204 NOLAAEGGSNVNHLIKVTOOSTEGYDSEGVYKAHDANLIYVTFEVDKVKASDIT 263
| : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : :
187 SNAKVEGTGTVTSKYVVEIGSI-EGHNNTNKV-EPPHAGRAVLKYLKEENGLHGGDFD 244
| : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : :
264 VIDKKTAVSDLTDSFTPIKIDNS-----GELIAGTYDNKNKOTTFYFDYDKENI 318
| : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : :
245 FTLSNVNVNTHGVTAKAYEIKNGSVYMATGEVLEGG-----KRIYTFNDIEKXUV 297
| : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : :
319 KAHKLITSTIDSKVPPNNNTKLDVEYKALSSVN---KITVEYQRPENRNTANLQSM 373
| : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : :
298 TAELEINLEFDPRVQTNNGN-----QITSTLENEQTSKELDVKKDGIGVYANLNGS 351
| : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : :
374 FTNIDKNTFTVEOTIYNLAKYSAKETNNVNSNGDEGSIIDDSIIIVYK-VGDNOL 432
| : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : :
352 IETFNKANNRFSHVAFIKP--NNGKTTSVYTGTLMKGSNONGNDPKVRIFFELGNNEDI 409
| : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : :
433 PDSNRIV---DYSEYEDVTNDYAOIG--NNNDVININGNDSPYIIIVISKYDKNKD- 485
| : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : :
410 AKS--VYANTTDTSKREYTSNMSGNLNQNNGSYSLNTEINDKYIVVHYDGEYLANGTDE 467
| : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : :
486 -DYTTIQOVTVMQTTINEYGEFRFASVDNTIAFSTSSGOGGDUPEPKTYKIGDYMED 544
| : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : :
468 VDFRT-QMVGHPQLKYYRGRYTLTMDNGVLVYENKANGN----- 508
| : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : :
545 VKDGS--IONTNDNEKPLSNVLVTLTPGTSKSVRTDDDKYQFQGYVD 593
| : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : :
509 -EKNGFIION-NKEFYKEDTITKETT--GOYDNKLVTVVEEYDSDTLID 555
| : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 2
PI4K.DICDI STANDARD: PRT; 1093 AA.
AC P54677;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE phosphatidylinositol 4-kinase (EC 2.7.1.67) (PI4-kinase) (PtdIns-4-
kinase) (PI4K-alpha).
GN PINK OR PI4K.
OS Dictyostelium discoideum (slime mold).
OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.

```

```

OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX3;
RX MEDLINE=96009592; PubMed=7565716;
RA Zhou K., Takegawa K., Emir S.D., Firtel R.A.;
RT "A phosphatidylinositol (PI) kinase gene family in Dictyostelium
discoidum: biological roles of putative mammalian p110 and yeast
Vps34p PI 3-kinase homologs during growth and development.";
RL Mol. Cell. Biol. 15:5645-5656(1995).
CC -I- FUNCTION: ACTS ON PHOSPHATIDYLINOSITOL (PI) IN THE FIRST
COMMITTED STEP IN THE PRODUCTION OF THE SECOND MESSENGER
INOSITOL-1,4,5--TRISPHOSPHATE.
CC -I- CATALYTIC ACTIVITY: ATP + 1-phosphatidyl-ID-myo-Inositol - ADP +
1-phosphatidyl-ID-myo-Inositol 4-phosphate.
CC -I- SIMILARITY: BELONGS TO THE PI3/PI4-KINASES FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL: U23479; AAA85725.1;
CC Dictydb: DD01102; PI4K.
CC InterPro: IPR00403; PI3_P14_kinase.
CC Pfam: PF00454; PI3_P14_kinase; 1.
CC SMART: SM00146; PI3K; 1.
CC PROSITE: PS00915; PI3_4_KINASE_1; 1.
CC PROSITE: PS00915; PI3_4_KINASE_2; 1.
CC PROSITE: PS00915; PI3_4_KINASE_3; 1.
CC TRANSFERASE; Kinase; Multigene family.
CC FT DOMAIN 146 156 ASP-RICH.
CC FT DOMAIN 184 200 POLY-ASN.
CC FT DOMAIN 203 206 POLY-ASN.
CC FT DOMAIN 227 233 POLY-ASN.
CC FT DOMAIN 277 315 POLY-ASN.
CC FT DOMAIN 441 445 POLY-ASP.
CC FT DOMAIN 455 466 POLY-THR.
CC FT DOMAIN 494 501 POLY-GLY.
CC FT DOMAIN 690 696 POLY-THR.
CC FT DOMAIN 751 754 POLY-GLN.
CC FT DOMAIN 761 772 POLY-THR.
CC FT DOMAIN 775 785 POLY-THR.
CC FT DOMAIN 833 1093 PI3K/PI4K.
SQ SEQUENCE 1093 AA; 123017 MW; 8762BC78355A635 CRC64;

Query Match          6.9%; Score 214; DB 1; Length 1093;
Best Local Similarity 20.2%; Pred. No. 0.028;
Matches 121; Conservative 101; Mismatches 224; Indels 152; Gaps 26;

1 HHHH--HHPSSDEK-----NDVINNOSINTD-DNNOIITKEETNNYDIERSD 49
| : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : :
167 HHYEDHDEHNDPKKDINSNNNNNNNNNNNNNNNDNNNEILPENSNSNTINDENQNG 226
| : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : :
50 RTESTTND-ENEATFLQKTPDNTHLTFEEYKSSSVSSSSSI--DTAQP----- 99
| : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : :
227 NSNNNNNIGENNNIKIDINSON-----KSDNIETLNTSLCEETKTSPIKDMEN 277
| : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : :
100 ----SHTTINREESVOTSDNVEDSHVDFANSKIKESN-----TESGKEENT 143
| : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : :
278 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 334
| : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : :
144 EOPNKVEDSTTSQPSG---YTNIDEKISNOELLNLPINEXENKARPLSTSAOPSIKR 200
| : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : :
335 SQPNDEIENITILKRNRIYKVEKRELARLRFCFMSYITGSRPLITPRRSSLPS 394
| : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : :
201 VTNOLAAEGGSNVNHLIKVTOOSTEGYDSEGVYKAHDANLIYVTFEVDKVKASD 260
| : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : :
395 PLISVSGKIGGNVYHILSPSSVDSTSLISEDKITKEEDENVVD---DDDEEVNSD 451
| : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : :

```



```
QY 261 -----TWTVIDIKNTVPSDLTDSFTIKIDNSG---EIIATGTYDNKKOITFTFDYV 312
D 452 FIPATATATTTTTPNHLSTKTSVGINSSTPININSAGAGGGEIHHI-----506
QY 313 DKENIKAHKLKLSYIDKSVNNNTKL---DVEYKTA-----LSSVKTITVE 358
D 507 -GTDDI-----SYLDCKCTPPAESKLSHDHFEFLSKHRCDIYLDLISFICKLAHIS 558
QY 359 ---YORENERTANLQSMFTNIDTKNHTVEQTIYINPLRSAKETVNVISGNGDEGSTII 415
D 559 KILPPIIDLRQAKLKH---EISLNTINPLGLV-PLMGSS-----596
QY 416 DDSITIVYVGNQNLPPDSNRITYSE-----YEDYTDNDYACLGNNNDVYNIFGNID 469
D 597 NHHCVR-----PREEVKILNSRERVPFLVLEVEISEHEHLSN-----637
QY 470 SPYIKYISKYDPKDDYTTIQOTVMTQNTINEYGE-----FRTASYDNTIARSTSS 522
D 638 ---LFEVYSSY---LQYTTGNSALAKKDDIKRKYSEKFKSFLNLSINSTISSSDS 688

SUIT 3
ARP_PLAFA STANDARD; PRT; 537 AA.
ID ARP_PLAFA
AC P04931;
DT 13-AUG-1987 (Rel. 05, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Asparagine-rich protein (Ag319) (ARP) (Fragmet).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
ON NCBI_TaxID=5833;
RX MEDLINE=66176787; Pubmed=2421257;
RA Stahl H.-D., Bianco A.E., Crewther P.E., Burkot T., Coppel R.L.,
RA Brown G.V., Anders R.F., Kemp D.J.;
RT "An asparagine-rich protein from blood stages of Plasmodium
RL falciparum shares determinants with sporozoites.";
CC Nucleic Acids Res. 14:3089-3102(1986).
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
RA EMBL; M24328; AAA29491.1; -
RA PIR; A23770; A23770.
RA Malaria.
FT NON_TER 1 1
FT NON_TER 537 537
SO SEQUENCE 537 AA; 63000 MW; A6D7B15318CC239 CRC64;

Query Match 6.6%; Score 203; DB 1; Length 537;
Best Local Similarity 17.8%; Pred. No. 0.035;
Matches 93; Conservative 106; Mismatches 212; Indels 112; Gaps 19;
```

```
QY 186 ARPLSTSAQPSIKRYVYNQLAAGQSNVHLLIVTQDSITEGVDSEGVKADAEHLI 245
D 164 AKNFYTN-----YNNKNNILKFTQNNNDNNINIEDNNNNNNNNNNNGVFSNTQNNMN 217
QY 246 YDVFEEVDKVKSGDITVDDIKNTVPSDLTDSFTIKIDNSGEIATGYDNKKQIT 305
D 218 RNNISINIKRLN-----NNNNNNNNKMSQDKQNSNNNFPMNTYQRRKSSMN 268
QY 306 YTFDYVDKYENIKAHKLKLSYIDKSVNNNTKLDVEYKTAALSSV-----352
D 269 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNSLSDMSPNYHAH 328
QY 353 -KTITVEYORENERTANLQSM-----FTYIDKNTV-----EDTIYIN-----391
D 329 VKMSMNVN-NESNTANPNQMFQGTNDNMKKRENNNNNYGYDDTVHNNNTPTDF 387
QY 392 -----PLRSKETEYVNI-----SGND-----EGSTIIDSTIIKYK-----VG 427
D 388 FSRAGVYNNNTLNNNNNNNSAVNNSSNGNNMKKNSGKAVADNNDNLNNKNNNNIN 447
QY 428 DNQNLPPDSNRITYSEY-----EDVTNDYACLGNNNDVIN 464
D 448 MNESINNNNTLNNNNNEYNQNNNEDDDDDWGEIGEDKYIDIN 490

RESULT 4
HMWL_MYCGE STANDARD; PRT; 1139 AA.
ID HMWL_MYCGE
AC 049413; 049365;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-Oct-2001 (Rel. 40, Last annotation update)
DE Cytoadherence high molecular weight protein 1 (Cytoadherence accessory
DE protein 1).
GN HMW1 OR MG312.
OS Mycoplasma genitalium.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2097;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33530 / G-37;
RX MEDLINE=96026346; Pubmed=756993;
RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,
RA Fleischman R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,
RA Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhmann J.L.,
RA Nguyen D.T., Uterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,
RA Tomb J.-F., Dougherty B.A., Bott K.F., Hu P.-C., Lucier T.S.,
RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;
RT "The minimal gene complement of Mycoplasma genitalium.";
RL Science 270:397-403(1995).
RN [2]
RP SEQUENCE OF 721-847 FROM N.A.
RC STRAIN=ATCC 33530 / G-37;
RX MEDLINE=94075230; Pubmed=8253680;
RA Peterson S.N., Hu P.-C., Bott K.F., Hutchison C.A. III;
RT "A survey of the Mycoplasma genitalium genome by using random
RT sequencing.";
RL J. Bacteriol. 175:7918-7930(1993).
CC -1- FUNCTION: COMPONENT OF THE CYTOSKELETON-LIKE STRUCTURE WHICH
CC STABILIZES THE SHAPE OF THE WALL-LESS MYCOPLASMA. THIS
CC CYTOSKELETON-LIKE NETWORK OF ACCESSORY PROTEINS CONTAINING HMW
CC PROTEINS 1 TO 5 ALLOWS THE PROPER ANCHORING OF CYTANELLE PROTEINS
CC IN THE MYCOPLASMA MEMBRANE AT THE ATTACHMENT ORGANELLE (BY
CC SIMILARITY).
CC -1- SUBCELLULAR LOCATION: LOCALIZES SPECIFICALLY TO THE ATTACHMENT
CC MEMBRANE (BY SIMILARITY).
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
```

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch)
CC -----
DR EMBL: U39712; AAC71534.1; -
DR EMBL: D02261; AAD12527.1; -
DR TIGR: MG312; -
CYCbaseReference: Structural protein: Complete proteome.
Q0 SEQUENCE 1139 AA; 130531 MW; 00111D3286C3DD856 CRC64;

Query Match	6.5%	Score 199.5	DB 1	Length 1139
Best Local Similarity	21.0%	Pred. No. 0.12		
Matches 129	Conservative 98	Mismatches 225	Indels 161	Gaps 27

OY 1 HHHHHHSSDEEKKDDVNNNO-SITFDNNNOI-----IKKEETNNYDG-----IE 44
 Db 578 HHEELKPAVEBQNNNVQVDFQVQANLDDNNEEIQPTAEKEVTTDFESSQAQVVDYQLPID 637
 45 KRSEDT-----ESTVNDENEAFTLOKTDQDNTHLLEEYKES-----SVESN--- 90
 Db 638 TDQDDQTFSSSEFTQPTVEEQFDQVNSVNDQFKPEITKEPELVESFNKODVNSNTN 697
 OY 91 --SSIDTAQPSHTTINNEESVQTSNDVEDSHVSPFANSKIKES-----NTEGKPENTIE 144
 Db 698 NLQKFDIOSDKKITTTKSSPQIPTLPISFVSNRIEYKVEYETALDNKEGQOQIT- 756
 OY 145 QPNKVKEDSTTSQPSGYTNIDE--KISNO---DELLNLPINEYENKARPSTTSAPDSI 198
 Db 757 --NSTTEDSKTIARTLSVOLQOINSLNNOISYTSSEVRLDKKDDQTLTNTVASEDQCPKI 814
 OY 199 KRVYNNOLAAGSGSVNHLIKYTDOSITEGYDDSGVTKARDAENLIYDVFEEVDYKYS 258
 Db 815 EVFKAKAPEVEHSTQTKQKQSEDEKSELDNNKKDL-----YKISELKR 860
 OY 259 GD-TWTVVDIDKNVYPSDLTDSFTLPKIDNGSEIIATGYDNKKQITVYTFDYDKYEN 317
 Db 861 GELNPTINFD-----ALFQNDYQMSVQKQSEFIHLNDFYTNKN 898
 OY 318 -----IKRAHL--KLTSYIDRSK--VPNNNTK-----LDVEYKTA 347
 Db 899 QISEKYLKKELOSELSRLIDQNNLNVAQNRNAKNLTLLQKEHIRSLASFAIAYKPS 958
 OY 348 -----LSSVNTTIVEYQRPNEKNTANLQSMFTNIDYKNTHTVEQTI-----Y 389
 Db 959 NSYEOLQKSGEIMRHVQCAIT-----ENEEKIESIQGLKOLKITYNSCCEITIMNNIKL 1013
 390 INPLYSKAKETNNVNSGND-----EGSTIIDDSITLIKVKYKGGDNQNLDSRIYDY 441
 Db 1014 DNTLRFACKEKEDPLLSNFDSTVTDNGLVEPQMLMD--LIDFSNFDN---ISNEQLDD 1067
 OY 442 SEYEDV-TNDYAYQALGNNDVYNINFGNIDSPIYIKVISKYDPNKDDYTIQOTV----- 495
 Db 1068 FIYEMADENIIFEEFEGFNND---FVDIDAK-VMSQMSAFSVNDLDIEFLVDRPNSFS 1122
 OY 496 -MGTINNEYTGEF 507
 Db 1123 LIDEDLFESSGDF 1135

 RESULT 5
 YMJL_CAEEL STANDARD: PRT: 918 AA.
 AC P34487;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Hypothetical 96.6 kDa protein F59B2.12 in chromosome III.
 GN F59B2.12.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdilitida; Rhabditoidea;
 OC Rhabdilitidae; Peloderinae; Caenorhabditis.
 CX NCBI_TaxID=6239;

[1]
RN SEQUENCE FROM N.A.
RP STRAIN-BRISTOL NZ;
RC MEDLINK=94150718; PubMed=7906398;
RX Wilson R., Alnsough R., Anderson K.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,
RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
RA Johnson L., Jones M., Kershaw J., Kirsten J., Laister N.,
RA Laitelle P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R.,
RA Sims M., Smaildon N., Smith A., Smith M., Sonhammer E., Staden R.,
RA Sulston J., Thierley-Mieg J., Thomas K., Vaudin M., Vaughan K.,
RA Waterston R., Watson A., Weinstock L., Wilkinson-Sproat J.,
RA Woldman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@sib-ch).

DR EMBL: Z11505; CAA77581.1; -.
DR PIR: S31132; S31132.
DR Wormpep: F5982.12; CE01024.
KW Hypothetical protein.
SQ SEQUENCE 918 AA; 96560 MW; E646FD8614945DE CRC64;

Query Match	6.48	Score 197	DB 1	Length 918
Best Local Similarity	20.58	Pred. No. 0.12	282	
Matches 142	Conservative 91	Mismatches	Indels 176	Gaps 27

```

0Y 2 HHHHH-----DEKKNVJNNNSINPDDNNQIIK----- 33
    || ||
Db 74 HKGHQSSSSSNTSLVYGADGNITENSEKKDGYKESKVDANENTIKSADGYI 133
0Y 34 -----KEETNNYDGIKRSERDTE-----STTNVDENEA-----TFLOKT 68
    :: ::
Db 134 ETGKSHNKSDDASVY-GLEKSSKTYADNGTMLSSNTNKKINNQSSAALDEGNEFVNOQ 192
0Y 69 PQDNH-----TEEVKSSSVESSENSIDTAQOPSHTTINREEVQYSDN--VEDS 119
    :: ::
Db 193 NADGFLRNNGHKNKTDELSHNVLDENAKMSIG-ADGSHITNRKKGIVGGSNHAASDA 251
0Y 120 HVSDFANSKIKESNTESGEEENTIEQPNKVKEDSTTSPGCTYNIDEKISNDELLNPI 179
    :: ::
Db 252 H-----SNPESIDAOGNKKKSQVSKKAASAGS--SNMDFE-SNLESJLNMDG 295
0Y 180 NEYENKARPLTTS-----AOPSIKRYVN-----QLAAEGSNVNHILKTYDSTIEG 228
    :: ::
Db 296 TSMASNTGNFNNTSYDKATAEEVMKKKNVNNADGTSMEASHAGSSNSKINSAGSSDLS 355
0Y 229 YDSEGVIAKHAENALIYDVFYEVDDKYVSGDTMYVIDDK----- 268
    :: ::
Db 356 MYGPRG-IKSHSTUKTGN--YALDEANGSASISEQIGKNQGRSLNESSIESGRKASR 412
0Y 269 NTVPEDLIDSTIPIKIDNSGEIINTGYIDNKKNKOITTYTIDYVUKYENIKAHKLSTYI 328
    :: ::
Db 413 NNTADTLDSDVANGTVSSSHKSSASGTSLDENHKKHTHALQASVDEHGMMKHSIDGSR 472
0Y 329 DKSVPNNNTKLDVEYKTA--LSASVN-----KTIYEVQRPENRFTANJQSMTYND 378
    :: ::
Db 473 NKKTEEFGESEMSASIKNADGTMQVYVKNDRNRTYKESKALBKNEKNSDGTGFXD-- 530
0Y 379 TKNHTVEQTIYINPLRYSAKETNVAISGNGEGSFTIIDDSTIIKYYKVGDNOLPDSNRI 438
    :: ::
Db 531 -----ESKGSNSRYVN--RTDGSNLAQVSGVSGKGVGASNETTIASSN-A 571

```

OY 439 YDYSEVEDYNDYAQGLNNNDVNFNIDSPYIKYISKYDPKDDYTTIQFTVMTQ 498
 DB 572 FNTSDAES--NOFDHLHKTANGTETTHAK--DS-----KOVAASANKSSIDTSMASVADAKG 625
 OY 499 TINEYTGFRASVYDNTAFSTSSGOGGDLPEPEKTYIGDYVMDVXKDJQNTNDMEK 558
 DB 626 NKRYVTSSQADSHD--AISASS-----DVDAKTYKHA--DNSE 660
 OY 559 PLSNVLTVPDGTSGKSVRTDEDKGYQFDG 589
 DB 661 SINDSSNOTASEHNDSSKQSEHEKRONADG 691
 RESULT 6
 BAF1_YEAST STANDARD; PRT: 731 AA.
 AC PL164:
 DT 01-JAN-1980 (Rel. 13, Created)
 DT 01-NOV-1980 (Rel. 16, Last sequence update)
 DT 01-FEB-1995 (Rel. 31, Last annotation update)
 DT Transcription factor BAF1 (ARS binding factor 1) (Protein ABF1)
 (Bidirectionally acting factor) (SFB-B) (DNA replication enhancer-
 binding protein OBF1)
 GN BAF1 OR ABF1 OR OBF1 OR YKLI12M OR YKL505.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OS Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OS Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OC NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A. PubMed-2511628;
 RX MEDLINE=90069539; PubMed-2511628;
 RA Diffley J.F.X., Stillman B.;
 RT "Similarity between the transcriptional silencer binding proteins
 ABF1 and RAP1.";
 RL Science 246:1034-1038(1989).
 RN [2]
 RP SEQUENCE FROM N.A. PubMed-2620828;
 RX MEDLINE=90152338; PubMed-2620828;
 RA Rhode P.R., Sweder K.S., Oegema K.F., Campbell J.L.;
 RT "The gene encoding ARS-binding factor I is essential for the
 RT viability of yeast.";
 RL Genes Dev. 3:1926-1939(1989).
 RN [3]
 RP SEQUENCE FROM N.A. AND SEQUENCE OF 128-154 AND 535-555.
 RX MEDLINE=90076149; PubMed-2686983;
 RA Halter H., Kavey B., Vandekerckhove J., Kiefer F., Gallwitz D.;
 RT "Sequence, expression and mutational analysis of BAF1, a
 RT transcriptional activator and ARS1-binding protein of the yeast
 Saccharomyces cerevisiae.";
 RL EMBO J. 8:4265-4272(1989).
 RN [4]
 RP SEQUENCE FROM N.A. AND PHOSPHORYLATION.
 RX MEDLINE=91239512; PubMed-2034654;
 RA Francesconi S.C., Eisenberg S.;
 RT "The multifunctional protein OBF1 is phosphorylated at serine and
 RT threonine residues in Saccharomyces cerevisiae.";
 RL Proc. Natl. Acad. Sci. U.S.A. 88:4089-4093(1991).
 RN [5]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92221689; PubMed-1561835;
 RA Jacquier A., Legrain P., Dujon B.;
 RT "Sequence of a 10.7 kb segment of yeast chromosome XI identifies the
 RT APN1 and the BAF1 loci and reveals one tRNA gene and several new open
 RT reading frames including homologs to RAD2 and kinases.";
 RL Yeast 8:121-132(1992).
 RN [6]
 RP SEQUENCE OF 585-731 FROM N.A.
 RX MEDLINE=94344141; PubMed-8065362;
 RA Butler A.R., White J.H., Fojlawiyo Y., Edlin A., Gardiner D.,
 RA Stark M.J.R.;
 RT "Two Saccharomyces cerevisiae genes which control sensitivity to G1
 RT arrest induced by Kluyveromyces fragilis toxin.";

RL Mol. Cell. Biol. 14:6306-6316(1994).
 CC -1- FUNCTION: TRANS-ACTING FACTOR IN THE REGULATION OF TRANSCRIPTION
 CC AND IN DNA REPLICATION. INVOLVED IN THE TRANSCRIPTION ACTIVATION
 CC OF A SUBSET OF RIBOSOMAL PROTEIN GENES. BINDS THE ARS-ELEMENTS
 CC FOUND IN MANY PROMOTERS. BINDS TO THE SEQUENCE 5'-TCN(7)ACG-3'.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- EXTENSIVELY PHOSPHORYLATED ON SER AND THR RESIDUES.
 CC -1- SIMILARITY: STRONG TO KLUYVEROMYCES ABF1, AND LOCAL TO YEAST RAP1.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sdb.ch/announce/>
 CC or send an email to license@sdb-sdb.ch).
 CC
 DR EMBL: X16385; CA344421.1; -
 DR EMBL: X51654; CA35966.1; -
 DR EMBL: M29067; AA66311.1; -
 DR EMBL: M63578; AA34823.1; -
 DR EMBL: S93804; AA22002.1; -
 DR EMBL: 228111; CA81951.1; -
 DR EMBL: X77511; CA54647.1; -
 DR PIR: S29870; S29870.
 DR TRANSFAC: T00056; -
 DR SGD: S0001595; ABF1.
 DR Transcription regulation; Activator; DNA-binding; Nuclear protein;
 KW Phosphorylation; Zinc-finger; Metal-binding; Zinc; DNA replication;
 KW Trans-acting factor.
 KM
 FT ZN_FING 49 71
 FT MOD_RES 624 624 PHOSPHORYLATION (BY PKC) (POTENTIAL).
 FT MUTAGEN 57 57 H->Q: LOSS OF DNA-BINDING.
 FT FT 71 71 C->S: LOSS OF DNA-BINDING.
 FT CONFLICT 125 125 N->K (IN REF. 3).
 FT CONFLICT 128 128 A->V (IN REF. 3 AND 5).
 FT CONFLICT 148 148 T->T (IN REF. 3 AND 5).
 FT CONFLICT 279 280 TN->NT (IN REF. 3 AND 5).
 FT CONFLICT 690 690 N->T (IN REF. 3, 5 AND 6).
 FT SEQUENCE 731 AA; 81749 MW; BF200372F3A9EA3F CRC64;
 SQ
 Query Match 6.3%; Score 195; DB 1; Length 721;
 Best Local Similarity 20.7%; Pred. No. 0.11; Indels 192; Gaps 33;
 Matches 138; Conservative 95; Mismatches 242;
 OY 1 HHHHSSDEKNDVNNOSIN-----TD--NNQIKKEETNNYDIEKR 46
 DB 105 HHHHNNNNNNDND--NNNGSNKASNDKLDPEVTDLEHLANIHPDTN--DKVESR 160
 OY 47 SEDRTSTTNVDENETFLQKTPQDNTHLTEEVKSSSVESNSSIDPAQPSHT--- 103
 DB 161 SNE--VNGNNDODADANNIFK--QGGVTIKNDIEDSINKASIDRLDDESGTGHGNDG 216
 OY 104 ---INREESVOT-----SDNVEDSHVS--DFANSKIKESVTESGREENTIEOPNKKEDS 153
 DB 217 NHRHNEEDDVHTQNTKNSDYVDEDIYVAIAVAANDSQSNKKDG-----KDD 268
 OY 154 TTSQPSGYTNIDEKISNODELNLPINEYENKARPLSTSAOPSIRYVNOIAAOGSN 213
 DB 269 ATNNNGGQDN---TNNDHNNNSNIN--NNNGSHGISHSPTSIRDSKNL----- 315
 OY 214 VNHLLKVTQSTIEGYDSGEVIAKAHAENLIYDFEVNDKYSQDPTMTVIDKNTVPS 273
 DB 316 -----DVFNASATDIDPGPEVYTKIEPY--HSHPLEDNLSTLKF 352
 OY 274 DLTFSTPIKIKNSGEI--IATGYDNKNKQIT--YTFDYDYK-----YE 316
 DB 353 -LT---KIPRILONDLKFDQILLESSTNNSNHTVSKFVSHYVESGLDLTKMORYLTAE 408
 OY 317 NIKAHV-----KLTSY-----IDRSKVPNNNTKLDEYKFTALSSVN--K 353
 DB 409 DFEKRLLSQARITTYIKARFVLLKKKMGREYNLDQSSSSNNNNNNNDGE---LSGTNLR 464

```

OY 354 TITVEQPNRPANLQSMETNIDKHHTEQFIYIPLRYSAETAVNISGNDGEST 413
Db 465 SNSIDYAKHQSISAGTSSN-----TTKN-----VANKKNDSDNNKNNNDAS 509
OY 414 IIDDSTIIKY---YKVGDNOLPDSNRIRYDSEEDVTNDYQOL-----GNNDVN 462
Db 510 NIMSEVLDKTSHTRYQ--PKMPSVKN--WSKPDQITHSDFVSWGIDESNDGNEHYH 563
OY 463 INFGNID-----SPYIIKYSKY-----DPNKDDYITIOQ?YIMOTTINEYGEFETA 510
Db 564 PTLAEVPAQAEARETAQLAIDKINSYKSIDKNGDGNHNSRHVYDENI,INDMSEDAHK 623
OY 511 SYDNTIAFSTSSGGGCGG-LLPEKTYRK-----DYWEDVDKDG 549
Db 624 SKRQHLSTITILEERNEDDKLPHEVYAEQURLLSHLKEVENDLCNNDDVDVMDVAVES 683
OY 550 IQNTNDN 556
Db 684 QYNKNTN 690

RESULT 7
LIPA_MYCPU ID LIPA_MYCPU STANDARD: PRT: 578 AA.
AC 050274; 098039;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Lipoprotein A precursor.
LIPA OR MYPU_5300.
OS Mycoplasma pulmonis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmales; Mycoplasma.
OX NCBI_TaxID:2107;

[1]
RN SEQUENCE FROM N.A.
RP STRAIN-KD735-15;
RX MEDLINE-20245550; PubMed-10781561;
RA Shen X., Gummlak J., Yu H., French C.T., Zou N., Dydvig K.;
RT "Gene rearrangements in the vra locus of Mycoplasma pulmonis.";
RL J. Bacteriol. 182:2900-2908 (2000).

[2]
RN SEQUENCE FROM N.A.
RP STRAIN-UAB C1P;
RX MEDLINE-21267165; PubMed-11353084;
RA Chambaud I., Hellig R., Ferris S., Barde V., Samson D., Gallsen F.,
RT Moszer I., Dydvig K., Wroblewski H., Viari A., Kocha E.P.C.,
RL Blanchard A.;
RT "The complete genome sequence of the murine respiratory pathogen
RL Mycoplasma pulmonis.";
RL Nucleic Acids Res. 29:2145-2153 (2001).

[3]
RN PARTIAL SEQUENCE FROM N.A.
RP STRAIN-KD735-15;
RX MEDLINE-96414471; PubMed-8817492;
RA Bhugra B., Voelker L.L., Zou N., Yu H., Dydvig K.;
RT "Mechanism of antigenic variation in Mycoplasma pulmonis: Interwoven,
RL site-specific DNA inversions.";
RL Mol. Microbiol. 18:703-714 (1995).

-1- SOBCELULOLAR LOCATION: Attached to the membrane by a lipid anchor.
(Potential).

-1- SIMILARITY: BELONGS TO THE PULMONIS LIPAB LIPOPROTEIN FAMILY.

-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on ways
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U023947; AAB41030.2; -.

```

[illegible]

OC Candida albicans (Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
 ON NCBI_TaxID=5476;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 10261;
 RX MEDLINE=96133936; PubMed=8552638;
 RA Gale C., Finkel D., Tao N., Meinke M., McClellan M., Olson J.,
 RA Gendrick K., Hostetter M.;
 RT "Cloning and expression of a gene encoding an integrin-like protein
 in *Candida albicans*.";
 RL Proc. Natl. Acad. Sci. U.S.A. 93:357-361(1996).
 CC -1- FUNCTION: COULD PLAY A ROLE IN ADHESION AND IN STE12-INDEPENDENT
 CC MORPHOGENESIS.
 CC -1- SUBCELLULAR LOCATION: CELL-SURFACE OF THE BLASTOSPORES.
 CC -1- SIMILARITY: SOME, WITH INTEGRIN ALPHA CHAINS.
 CC -1- SIMILARITY: CONTAINS 1 PH DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U35070; AAA96019.1;
 DR InterPro: IPR001849; PH.
 DR Pfam: PF00169; PH.1.
 DR SMART: SM00233; PH.1.
 DR PROSITE: PS50003; PH_DOMAIN; 1.
 DR Cell adhesion; glycoprotein.
 KW DOMAIN 1527 1636
 FT SITE 1149 1151 PH.
 FT CARBOHYD 50 50 CELL ATTACHMENT SITE (POTENTIAL).
 FT CARBOHYD 154 154 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 216 216 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 223 223 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 265 265 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 268 268 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 285 285 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 289 289 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 306 306 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 314 314 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 446 446 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 583 583 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 638 638 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 670 670 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 691 691 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 817 817 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 838 838 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 841 841 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 1018 1018 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 1082 1082 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 1100 1100 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 1103 1103 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 1113 1113 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 1200 1200 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 1571 1571 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 1593 1593 N-LINKED (GLCNAC...) (POTENTIAL).
 FT DOMAIN 95 106 POLY-GLN.
 FT DOMAIN 283 286 POLY-ASN.
 FT DOMAIN 1283 1292 POLY-HIS.
 FT DOMAIN 1651 1660 POLY-GLN.
 SO SEQUENCE 1664 AA; 187859 MW; BA2EE0DDC8196790 CRC64;

Query Match 6.1%; Score 189.5; DB 1; Length 1664;
 Best Local Similarity 20.4%; Pred. No. 0.49;
 Matches 131; Conservative 105; Mismatches 236; Indels 171; Gaps 31;

15 DIVINNOS---INTDNNQIIRK-----EFNNYNGIIRKSDREESTT 55

Db 138 DKVENNHAPTIINTSPKSIKKATPKASPKKVAFTVTPNPEIHNP--DNKVEEDQSOQ 195
 QY 56 NVDEENATFLOKTPODNHLTEEEVKSSVESSNSIDTAQOQPSHTT-----INRE 108
 Db 196 KEDSEYEPPLIQ-----HOKKDPQFNVSDDEDTNVA5VPT-PLHTHTKPPFAQLNN 247
 QY 109 SVQTSQDNVEDSHVD-----FANSKIKE-----STESGKEBNTTEQPKKVEDSTT 155
 Db 248 EV---NSEPALTDMLKRENFNSLDEKYNILSTPTNNNSKRVSDMDSHO-----N 299
 QY 156 SQPSGYNIDKISQDDELNPIVEYENKARPL-STISAPSTIRKRVVNOAAGSNV 214
 Db 300 LODASKRNTNENHNLSPALAPKNDIEN--PLNSLTNADISLSSGSSQSOGLSLND 356
 QY 215 NHLIV---TDOSTEGYDSDGVKADAEVLVDYFEVDKVSQDFTVVDIKNTV 271
 Db 357 NRVLESVQSPKKNVPGSLNDGKGFSD-----EVYESLL 392
 QY 272 PSDLT-DSFTTPKIKD-----NSGEIATGYDNKNKOITYFTDYVDKENIKAH----- 321
 Db 393 PDLSDRLKLETKEDHDAPEHNNEFIDAKSTNKG-OLVSSDDHLSFDRSYNHTQOS 451
 QY 322 -LKLTSTYIDKSKVPMN-----NPKLDVYKTAALSSVVK 353
 Db 452 ILLNLNSASQSOISLNALEKORQOEOQAEPDEETSFSDNKKVQEPKSNLEFKV 511
 QY 354 TTVEYQRPNERNTANLSMFTNIDTKNH-TVEQTYINPLRYSAKETNVNISG----- 406
 Db 512 TIKKEPVASATELKARKRESSRLAIKNDELAEPADLHPKKNENANSHVEDTDLAKKA 571
 QY 407 -NGDEGSTIIDSTTIKY-YKVGDNQNLPSNRITYSEVEDYVNDYQAQANN--NDV 461
 Db 572 LNDDESDTTONSTKMSIRFHIDSDMKLEDSND--GDRED--NDDISREFKSDILND 625
 QY 462 N-----INGNDISPIYIKVSKYDKPKDYTTIQGYVMQTTINETGEFR-----T 509
 Db 626 SOTSIIIGDKYGNSSSEITTKTLA--PRSDNNKREKSKSLSDPANNE5LQQLVPHPT 682
 QY 510 ASYDNTIAFSTSSGOGCDLPPK-----TYKIGDY-VVEDVDK 547
 Db 683 KEDDSIIANSSNIA-----PPEELTLPEVEANDVS5SFNDYK 719
 RESULT 9
 EBA1_PLAFC STANDARD; PRT; 1435 AA.
 ID EBA1_PLAFC
 AC P19214;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 01-FEB-1996 (Rel. 33, Last annotation update)
 DE Erythrocyte-binding antigen EBA-175.
 OS Plasmodium falciparum (isolate Camp / Malaysia).
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.
 ON NCBI_TaxID=5835;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90377299; PubMed=2204835;
 RA Sim B.R.L.;
 RT "Sequence conservation of a functional domain of erythrocyte binding
 antigen 175 in *Plasmodium falciparum*.";
 RL Mol. Biochem. Parasitol. 41:293-296(1990).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X52524; CA36756.1;
 DR PIR: S11561; S11561.

KW Antigen. 159 1104 ESSENTIAL FOR BINDING TO
 FT DOMAIN
 FT VARIANT 1031 1031 ERYTHROCYTES.
 E -> V (IN STRAINS FCR-3 AND ITG).
 SQ SEQUENCE 1435 AA; 167389 MW; 32AA309021B1C3D6 CRC64;
 Query Match 6.1%; Score 188.5; DB 1; Length 1435;
 Best Local Similarity 20.5%; Pred. No. 0.45;
 Matches 114; Conservative 102; Mismatches 199; Indels 141; Gaps 28;
 QY 8 SSBREKNDVYNNNSINTDNNQIIEKKEETNNVDG-----IEKRSDEPTSTNTVDEN 60
 DB 875 NSDESEETVYNN-----ISKSPINNGDSSGSSATVSESSNTGLSDIDDRN 923
 QY 61 EATFLQKTPQDTHLLEEV--KESSSVSSNSIDTAQPSHTTINREESVQTSQNVBD 118
 DB 924 GDTFVATQDTANT---EDVIRKENADKDEDEKGADEERHSTSESLSPSEKMLTDNEG 979
 DB 119 SHVSPANKIKE--SNTESSGKEENTIEQPNVKE-DSTTSQPSGTYNDE----- 166
 DB 980 NSLN--HEVEKEHTNSDNVQSGIVMNVKEKLTDLNPS--SSLDKGAHEELSE 1034
 QY 167 -KISNDELINLPINEYENKARPLSTTSAQPSIKRYTVNQLAEOGSNNHLIKYTDQSI 225
 DB 1035 PMLSSQDMSNTP-----GPLDNTS-EETTERISNNEYKVN-----REDERTL 1077
 QY 226 TEGYDDSEGVIAH-----DAENLIYDTFEVDKVKSGDTMTVDIDKNTPVSDLTDSF 279
 DB 1078 TREYEDI--VLKSHMRSEDGE--LYD-----ENSDLSIVNDESEDAEMAKKND 1124
 QY 280 TTPKIDNNGEILANTCTYNNKKQITFTDYVDKYENKALKLTSTIDSKSVPNNTK 339
 DB 1125 T-SESHNSQHSIESQONDKTYGDLGTHVQN---EISVPYGEIDEXLRSEKSK 1179
 QY 340 LDVEYKALTSVNTKTTFVEYQRPENRNTANLQSM--FTNIDKNHTVEQITINPLRYSA 397
 DB 1180 IKAKEERLSHND---IKINP-EDRNSNTLHLDIRBENRHLTNINISQERDQ 1234
 QY 398 K--ETNVNISGNGDEGSTIIDSTTIKYKVGDNONLPDSNRYDYSEEDVTNDVYO 454
 DB 1235 KQGFHTMNLHGDG-----VSRSQINSHSHGNROD-----R 1266
 QY 455 LGNNNDV-----NINEGNDSPYIKVSKYD-----ENKDDYTIQOTYMTQTTI 500
 DB 1267 GGNNGVWLMRSNNNNFNIPRY---NLTDKTLIDLLEYENNDSTI-KELIKLAEI 1320
 DB 501 NEYTGERTASYDNTI 516
 DB 1321 NKCENISVYCDHMI 1336
 RESULT 10
 LMA2_HUMAN STANDARD: PRT: 3110 AA.
 AC P24043; Q14736; Q93022;
 DT 01-MAR-1992 (Rel. 21, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Laminin alpha-2 chain precursor (Laminin M chain) (Merosin heavy chain)
 GN LAMA2 OR LAMM.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Plimates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Placenta;
 RX MEDLINE=94124633; PubMed=8294519;
 RA Vuolteenaho R., Nissinen M., Salin K., Byers M., Eddy R.,
 RA Hironen H., Shows T.B., Sariola H., Engvall E., Tryggvason K.,
 RT "Human laminin M chain (merosin): complete primary structure,"

RT chromosomal assignment, and expression of the M and A chain in human
 RT fetal tissues.";
 RL J. Cell Biol. 124:381-394(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97066955; PubMed=8910357;
 RA Zhang X., Vuolteenaho R., Tryggvason K.,
 RT "Structure of the human laminin alpha2-chain gene (LAMA2), which is
 RT affected in congenital muscular dystrophy.";
 RL J. Biol. Chem. 271:27664-27669(1996).
 RN [3]
 RP SEQUENCE OF 1981-3110 FROM N.A., AND PARTIAL SEQUENCE.
 RC TISSUE-Placenta;
 RX MEDLINE=90238994; PubMed=2185464;
 RA Ehrig K., Leivo I., Argraves W.S., Ruoslahti E., Engvall E.,
 RT "Merosin, a tissue-specific basement membrane protein, is a
 RT laminin-like protein.";
 RN Proc. Natl. Acad. Sci. U.S.A. 87:3264-3268(1990).
 RN [4]
 RP VARIANTS GLN-545, HIS-619, LEU-919, HIS-2586 AND LYS-2614.
 RA Panicker S.G., Mendell J.T., Chen L., Feng B., Sahenk Z.,
 RA Mazluf G.A., Amato A.A., Mendell J.R.,
 RT "Novel single base polymorphisms and rare sequence variants in
 RT the laminin 2-chain coding region detected by RNA/SSCP analysis.";
 RN Hum. Mutat. 13:174-174(1999).
 RN [5]
 RP ERRATUM.
 RA Panicker S.G., Mendell J.T., Chen L., Feng B., Sahenk Z.,
 RA Mazluf G.A., Amato A.A., Mendell J.R.,
 RL Hum. Mutat. 13:340-340(1999).
 CC - FUNCTION: BINDING TO CELLS VIA A HIGH AFFINITY RECEPTOR, LAMININ
 CC IS THOUGHT TO MEDATE THE ATTACHMENT, MIGRATION, & ORGANIZATION OF
 CC CELLS INTO TISSUES DURING EMBRYONIC DEVELOPMENT, & INTERACTING
 CC WITH OTHER EXTRACELLULAR MATRIX COMPONENTS.
 CC - SUBUNIT: LAMININ IS A COMPLEX GLYCOPROTEIN, CONSISTING OF THREE
 CC DIFFERENT POLYPEPTIDE CHAINS (ALPHA, BETA, GAMMA), WHICH ARE BOUND
 CC TO EACH OTHER BY DISULFIDE BONDS INTO A CROSS-SHAPED MOLECULE
 CC COMPRISING ONE LONG & THREE SHORT ARMS WITH GLOBULES AT EACH END.
 CC THE ALPHA-2 CHAIN IS A SUBUNIT OF LAMININ-2 (MEROSIN) AND LAMININ-
 CC 4 (S-MEROSIN).
 CC - SUBCELLULAR LOCATION: EXTRACELLULAR; FOUND IN THE BASEMENT
 CC MEMBRANES (MAJOR COMPONENT).
 CC - TISSUE SPECIFICITY: PLACENTA, STRIATED MUSCLE, PERIPHERAL NERVE,
 CC CARDIAC MUSCLE, PANCREAS, LUNG, SPLEEN, KIDNEY, ADRENAL GLAND,
 CC SKIN, TESTIS, MENINGES, CHOROIOID PLEXUS, AND SOME OTHER REGIONS OF
 CC THE BRAIN; NOT IN LIVER, THYMUS AND BONE.
 CC - DOMAIN: THE ALPHA-HELICAL DOMAINS I AND II ARE THOUGHT TO INTERACT
 CC WITH OTHER LAMININ CHAINS TO FORM A COILED COIL STRUCTURE.
 CC - DOMAIN: DOMAINS VI, IV AND G ARE GLOBULAR.
 CC - DISEASE: DEFECTS IN LAMA2 ARE THE CAUSE OF MEROSIN-DEFICIENT
 CC CONGENITAL MUSCULAR DYSTROPHY (MCMD).
 CC - SIMILARITY: CONTAINS 1 LAMININ N-TERMINAL DOMAIN (DOMAIN VI).
 CC - SIMILARITY: CONTAINS 2 LAMININ DOMAINS.
 CC - SIMILARITY: CONTAINS 5 LAMININ G-LIKE DOMAINS.
 CC - This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL, 226653; CAAB1394.1; -
 CC EMBL, 226653; CAAB1394.1; -
 CC EMBL, 066796; AAB18388.1; JOINED.
 CC EMBL, 066733; AAB18388.1; JOINED.
 CC EMBL, 066734; AAB18388.1; JOINED.
 CC EMBL, 066735; AAB18388.1; JOINED.
 CC EMBL, 066736; AAB18388.1; JOINED.
 CC EMBL, 066737; AAB18388.1; JOINED.
 CC EMBL, 066738; AAB18388.1; JOINED.
 CC EMBL, 066739; AAB18388.1; JOINED.

EMBL; U66740; AAB18388.1; JOINED.
DR EMBL; U66741; AAB18388.1; JOINED.
DR EMBL; U66742; AAB18388.1; JOINED.
DR EMBL; U66743; AAB18388.1; JOINED.
DR EMBL; U66745; AAB18388.1; JOINED.
DR EMBL; U66746; AAB18388.1; JOINED.
DR EMBL; U66747; AAB18388.1; JOINED.
DR EMBL; U66748; AAB18388.1; JOINED.
DR EMBL; U66749; AAB18388.1; JOINED.
DR EMBL; U66750; AAB18388.1; JOINED.
DR EMBL; U66751; AAB18388.1; JOINED.
DR EMBL; U66752; AAB18388.1; JOINED.
DR EMBL; U66753; AAB18388.1; JOINED.
DR EMBL; U66754; AAB18388.1; JOINED.
DR EMBL; U66755; AAB18388.1; JOINED.
DR EMBL; U66756; AAB18388.1; JOINED.
DR EMBL; U66757; AAB18388.1; JOINED.
DR EMBL; U66758; AAB18388.1; JOINED.
DR EMBL; U66759; AAB18388.1; JOINED.
DR EMBL; U66760; AAB18388.1; JOINED.
DR EMBL; U66761; AAB18388.1; JOINED.
DR EMBL; U66762; AAB18388.1; JOINED.
DR EMBL; U66763; AAB18388.1; JOINED.
DR EMBL; U66764; AAB18388.1; JOINED.
DR EMBL; U66765; AAB18388.1; JOINED.
DR EMBL; U66766; AAB18388.1; JOINED.
DR EMBL; U66768; AAB18388.1; JOINED.
DR EMBL; U66769; AAB18388.1; JOINED.
DR EMBL; U66770; AAB18388.1; JOINED.
DR EMBL; U66771; AAB18388.1; JOINED.
DR EMBL; U66772; AAB18388.1; JOINED.
DR EMBL; U66773; AAB18388.1; JOINED.
DR EMBL; U66774; AAB18388.1; JOINED.
DR EMBL; U66775; AAB18388.1; JOINED.
DR EMBL; U66776; AAB18388.1; JOINED.
DR EMBL; U66777; AAB18388.1; JOINED.
DR EMBL; U66778; AAB18388.1; JOINED.
DR EMBL; U66779; AAB18388.1; JOINED.
DR EMBL; U66780; AAB18388.1; JOINED.
DR EMBL; U66781; AAB18388.1; JOINED.
DR EMBL; U66782; AAB18388.1; JOINED.
DR EMBL; U66783; AAB18388.1; JOINED.
DR EMBL; U66784; AAB18388.1; JOINED.
DR EMBL; U66785; AAB18388.1; JOINED.
DR EMBL; U66786; AAB18388.1; JOINED.
DR EMBL; U66787; AAB18388.1; JOINED.
DR EMBL; U66788; AAB18388.1; JOINED.
DR EMBL; U66789; AAB18388.1; JOINED.
DR EMBL; U66790; AAB18388.1; JOINED.
DR EMBL; U66791; AAB18388.1; JOINED.
DR EMBL; U66792; AAB18388.1; JOINED.
DR EMBL; U66793; AAB18388.1; JOINED.
DR EMBL; U66794; AAB18388.1; JOINED.
DR EMBL; U66795; AAB18388.1; JOINED.
DR EMBL; M59892; AAB18388.1; JOINED.
DR PIR; A35899; M59892.
DR HSP; P02468; 1KLO.
DR MIM; 156225;
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR001866; LamNT.
DR InterPro; IPR000034; Laminin_B.
DR InterPro; IPR002049; Laminin_EGF.
DR InterPro; IPR001791; Laminin_G.
DR Pfam; PF00052; Laminin_B; 2.
DR Pfam; PF00053; Laminin_EGF; 15.
DR Pfam; PF00054; Laminin_G; 5.
DR Pfam; PF00055; Laminin_Nterm; 1.
DR PRINTS; PR00011; EGF_LAMININ.
DR ProDom; PD002082; LamNT; 1.
DR ProDom; PD003031; Laminin_B; 2.
DR SMART; SM00180; EGF_Lam; 14.
DR SMART; SM00001; EGF-like; 3.
DR SMART; SM00281; LamB; 2.

DR SMART; SM00282; LamG; 5.
DR SMART; SM00136; LamNT; 1.
DR PROSITE; PS00022; EGF_1; 11.
DR PROSITE; PS01186; EGF_2; 3.
DR PROSITE; PS01248; LAMININ_TYPE_EGF; 14.
DR PROSITE; PS50025; LAM_G_DOMAIN; 5.
DR GlycoProtein; Basement membrane; Extracellular matrix; Coiled coil;
KV Laminin EGF-like domain; Cell adhesion; Repeat; Signal; Polymorphism.
KW Laminin EGF-like domain; Cell adhesion; Repeat; Signal; Polymorphism.
FT SIGNAL 1 22
FT CHAIN 23 3110 LAMININ ALPHA-2 CHAIN
FT DOMAIN 23 286 LAMININ N-TERMINAL (DOMAIN VI).
FT 287 343 LAMININ EGF-LIKE 1.
FT DOMAIN 344 413 LAMININ EGF-LIKE 2.
FT DOMAIN 414 468 LAMININ EGF-LIKE 3.
FT DOMAIN 469 517 LAMININ EGF-LIKE 4.
FT DOMAIN 518 527 LAMININ EGF-LIKE 5 (N-TERMINAL).
FT DOMAIN 528 723 LAMININ DOMAIN IV 1 (DOMAIN IV B).
FT DOMAIN 724 756 LAMININ EGF-LIKE 5 (C-TERMINAL).
FT DOMAIN 757 806 LAMININ EGF-LIKE 6.
FT DOMAIN 807 864 LAMININ EGF-LIKE 7.
FT DOMAIN 865 917 LAMININ EGF-LIKE 8.
FT DOMAIN 918 966 LAMININ EGF-LIKE 9.
FT DOMAIN 967 1013 LAMININ EGF-LIKE 10.
FT DOMAIN 1014 1059 LAMININ EGF-LIKE 11.
FT DOMAIN 1060 1105 LAMININ EGF-LIKE 12.
FT DOMAIN 1106 1165 LAMININ EGF-LIKE 13.
FT DOMAIN 1166 1175 LAMININ EGF-LIKE 14 (N-TERMINAL).
FT DOMAIN 1176 1379 LAMININ DOMAIN IV 2 (DOMAIN IV A).
FT DOMAIN 1380 1419 LAMININ EGF-LIKE 14 (C-TERMINAL).
FT DOMAIN 1420 1468 LAMININ EGF-LIKE 15.
FT DOMAIN 1469 1526 LAMININ EGF-LIKE 16.
FT DOMAIN 1527 1573 LAMININ EGF-LIKE 17.
FT DOMAIN 1574 2144 LAMININ II AND I.
FT DOMAIN 2145 2328 LAMININ G-LIKE 1.
FT DOMAIN 2340 2521 LAMININ G-LIKE 2.

Query Match 6.0%; Score 186.5; DB 1; Length 3110;
Best Local Similarity 19.9%; Pred No. 1.3; Mismatches 215; Gaps 32;
Matches 142; Conservative 102; Mismatches 256; Indels 215; Gaps 32;

QY 28 NNOIKKEET-----NNTDGLKRESDRTSTNVNENAFLOQTPDNTLHTEE 78
DB 1693 NEKAIKNETLTGTRDEAFERNLEGLQK-----EIDQMKELRRKRLTEQREIAD 1742

QY 79 EYKSSSVSSNSSIFTPAQPSTHTIN---REDSVOTSDNVEDS-HVSDPANSKIESN- 133
DB 1743 ELVAEALTKKVKILFGESGEMEMEKDLREKLADYKKNVDAMDLLREATDKIREANR 1802

QY 134 -----TESGKE--ENTIEOPNKV-----KEDSTS 156
DB 1803 LRAVNOKNMTALEKKKEAVESGRJOIENTLKBENDILDEANRLADEINSIIVDEIOTK 1862

QY 157 OPSGYTINDEKISN---ODELLNPINEYENKARPLSTSA-----OFSIKRVTVNOL 206
DB 1863 LPPMSEELNDKIDDLDSQELKDRKLAEKVSQAESHAAQLNDSSAVLGLIDEAKNISFNAT 1922

QY 207 AAEQG-SVNHLLIKVTDOSTTSGYIDDSSEGYIKAHDEMNIIVDYFEYVDKVSQGDIMYD 265
DB 1923 AAFKAYSNIKDYI-----DEAEKVAK--EAKDLAHEAT-----KLAAGPR--G 1961

QY 266 IDKNTPSPDLTDSFTI-----PKIDNSGEIINTGT-----YDNKNQIYTFPDYVD 313
DB 1962 LKEDAKKGLQKSFRLNEKKKLANDYKNEEDHNLGKTRIEADARNQDLTLTMDTIG 2021

QY 314 KYENI--KAHLKLTXYIDRSKYPNNNTK-----LDVEYRTALSSVKKTTT 356
DB 2022 KLSAIFNDPAKKAQAVKDKARQANDTAKDYLAQITELHQNLQDLKKRYNKLADSVAKTA 2081

QY 357 VEYQRPENENR-TANLOSMTFNID-----TKNRTVEQTIYINPLKRS 396
DB 2082 V-VKDESKNKTIIADADATYVNLQEQADRDLIDKLPIKELEDNKKNISSEIKELINOARQO 2140

QY 397 AKETYNVISGND-----EGS-----TIIDSTIIKV-----YK 425


```

DB 2141 ANSIRKVSAGGDCIRTKYKPEIKKSYNNIYVAVKTAADVADNLLFYLSAKFIIDFLAEMR 2200
QY 426 VGDNQNLPPSNRIYDYSEEDVTND-----YAQLGNNDVNIINFGINDSPIYIKVIS 478
DB 2201 KGVSTSLAMPVSGVSGVGPDLIDDSYVRIYASRTGRGRTISVR--ALDGPKASIVPS 2258
QY 479 -KXDPKKDDYTTIQVTVMTQNTINEYTGERTASVYNTIAFSTSSGOGDLPPEKTYKI 537
DB 2259 THHSTSPGGTITIDVANNMLFVGLTGKLRKADAVRTTFTSCMETYFDNKP-----I 2313
QY 538 GBYVMDVDKDGIONTNDNEKPLSNVLTLYEPD--GTSKSVRT--DEDKYOPDG 589
DB 2314 G--LWMPREKES-----DCKGCTVSPQVDESEGTIOFDG 2345

RESULT 11
YBET_YEAST STANDARD: PRT: 1381 AA.
P34216;
01-FEB-1994 (Rel. 28, Created)
01-OCT-1994 (Rel. 30, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Hypothetical 150.8 kDa protein in SEC17-QCRL intergenic region.
YBL047C OR YBL0520 OR YBL0501.
Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_Taxid=4932;
RN [1]
RP SEQUENCE OF 1-961 FROM N.A.
RC STRAIN=S288C;
RX MEDLINE=95116707; PubMed=7871888;
RA De Wergifosse P., Jacques B., Joniaux J.-L., Purnelle B., Skala J.,
RA Goffeau A.,
RT "The sequence of a 22.4 kb DNA fragment from the left arm of yeast
RT chromosome II reveals homologues to bacterial proline synthetase and
RT murine alpha-daptin, as well as a new perase and a DNA-binding
RT protein."
RT Yeast 10:1489-1496(1994).
RL [2]
RP SEQUENCE OF 579-1381 FROM N.A.
RC STRAIN=S288C;
RX MEDLINE=94205266; PubMed=8154187;
RA Scherens B., el Bakhouy M., Vierendeels F., Dubois E., Messenguy F.,
RA "Sequencing and functional analysis of a 32,560 bp segment on the
RA left arm of yeast chromosome II. Identification of 26 open reading
RA frames, including the KIP1 and SEC17 genes."
RA Yeast 9:1355-1371(1993).
CC -1- SIMILARITY: BELONGS TO THE VDP/USO1/YBL047C FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: 235808; CAAB4867.1;
CC EMBL: X78214; CAAS5048.1;
CC EMBL: 223261; CAAB0797.1;
CC PIR: S45781; S45781.
CC SGD: S0000143; YBL047C.
CC InterPro: IPR000261; EFS15_repeat.
CC InterPro: IPR000449; UBA.
CC Pfam: PF00036; ehand; 2.
CC SMART: SM00627; UBA; 1.
CC SMART: SM00054; Eph; 2.
CC SMART: SM00027; EH; 3.
CC SMART: SM00165; UBA; 1.
CC Hypothetical protein.
CC SEQUENCE 1381 AA; 150783 MW; 626FD26IDCBA7D99 CRC64;

```

```

Query Match 6.0%; Score 186; DB 1; Length 1381;
Best Local Similarity 20.0%; Pred. No. 0.55;
Matches 143; Conservative 115; Mismatches 262; Indels 194; Gaps 34;

QY 7 PSSDEKNDYINN-NQINMTDNN--QIIKKE-----TNNYDIERKSEDRSTSTN 56
DB 437 PSPTKAQTVVQNNNTNNSFYDNNNGATLQQQQPQPPPLTTHSSSGIKFT-----PTSN 491
QY 57 VDENEATFLPKTPQDNTHTLEEE-----VKESSSVSESSNSSIDTAQPSHTTINR 106
DB 492 PQQ---SIKKEEEDQQLRESSDTSAPPPPKHASSPVKRTATTLTPQVNFVFSM 548
QY 107 -----EESVQTSQD-NVEDSH 120
DB 549 PAGAATSAATGAAGAAGAAALGASAFSSSSNNNAFKNODLFADGEASQLSNATTEMAN 608
QY 121 VSPFANSKIKESNTESGKEENTIEOPNKYKEDSTQSPSGYTNI-----DEKISNDEL-- 174
DB 609 LSNQVNSLSKQASITNDKSRATQELKRTYTEMKNSIQIK-LNNLRSTHDQNVACTOLEA 667
QY 175 LNLPIIN-EYENKARPLSTTSAPPSIKRVYVNOALABQGSNVHLIKVTDQSITEGYDSE 233
DB 668 QVLQVKNENETLAAQLAVSEA-----NYHAAE--SKINELTDLQESQTKNAELKE 716
QY 234 GYIKAHDAENLIYDTVEYVDKVKSGDTMTVDIDKNTVPSDLTDSFTIKRINSGEIIA 293
DB 717 QITNLNSMTASLQSLNEKQOQVKQERSM-VDVNSKQLELNQVTVANLQKEIDGLEKIS 775
QY 294 TCTYDNKNKQITVTFTDYVDKYENIKAKLTSYIDSKVPNNNTKLDEYKALSSVVK 353
DB 776 --VYLLKQKEL--NDQKYVEQHAQIQ-AKYQDLS---KNDTDL-IDREKQLEERNR 824
QY 354 TTFVEYQRPENENTANLQSMFTNIDYKNHTVEQT-----IYNPIAR-YSAKETNVN 403
DB 825 QIE-EOENLYHQVSKLQEMFDDLQSRKASFQADQELKERNIEYANNVRELSEERQMLA 883
QY 404 ISNGDEGSTII-----DQSTIIK-VYVVEGNNLPPSNRIYDIS 442
DB 884 MGQLPEDARDIILAKSASNTDTTKEATSRGNHEDTVSKFVETTVENSNL-NNRYKQDE 942
QY 443 E-----YEDVTNDYVQLGNNND--VINFG-----NIDSPYIIIVIS-KYDPNKDY- 487
DB 943 EKTERTESDVPFDVDTLQSGSDSENNATNNGSQSNEANPULTELTSLDFREGDLNKG 1002
QY 488 -----TTIQQVTVMTQNTINEYTGERTASVYNTIAFSTSSGOGDLP 531
DB 1003 IPRQSLTSSVANNAPQSVRDVLELPETLEERTINNTANRONTGNLSHIPGE----- 1055
QY 532 ETKYKIGDYVWEVDKDGIONTNDNEKPLSNVLTLYEPD--TSKSVRTDEDK 584
DB 1056 -----WEATPATASTDVLNSN-----TTVEIDEGSTTKRANSNDEG 1092

RESULT 12
YNP4_YEAST STANDARD: PRT: 1165 AA.
P53950;
01-OCT-1996 (Rel. 34, Created)
01-OCT-1996 (Rel. 34, Last sequence update)
01-OCT-1996 (Rel. 34, Last annotation update)
Hypothetical 128.1 kDa protein in OMP2-MG65 intergenic region.
YND054W OR N2467 OR YNL2467W.
OC Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_Taxid=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / FY1676;
RX MEDLINE=96021608; PubMed=8533472;
RA Bergez P., Dolignon F., Crouzet M.;

```


RT "the sequence of a 44 420 bp fragment located on the left arm of
 RT chromosome XIV from Saccharomyces cerevisiae."
 RL Yeast 11:967-974(1995).
 RN [2]
 RP ERRATUM.
 RX MEDLINE=97060022; Pubmed=8904343;
 RA Berger P, Dolgon F, Crouzet M.;
 RL Yeast 12:297-297(1996).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U12141; AAA99658.1; -
 DR EMBL: 271330; CA95925.1; -
 DR SCD: S0004999; YNL054W.
 DR Hypothetical protein.
 SEQID 1165 AA; 128140 MW; 0414978AA3066CE7 CRC64;

Query Match 6.0%; Score 185.5; DB 1; Length 1165;
 Best Local Similarity 18.6%; Pred. No. 0.47; Indels 175; Gaps 27;
 Matches 120; Conservative 102; Mismatches 249;

QY 4 HHHSSEDEKNDVNNOSINT-----DONNOIKKEETNNYDGIKESDRE----- 52
 DB 299 NNSAADNENNNKKKKNNKNSGNERNDPSKICTSTKAPSTAPLGSDDNVALTAS 358
 QY 53 -STTVNDEATFLQKTPQDTHLTETEVKSSSVESNSSID-TAQPSHTTINREESV 110
 DB 359 VSSSNAD-NHNNNNKKTSSNNNGNSASAKTNADIKNSADLSASTSNNAINDSHE 417
 QY 111 QTSN-----VEDSHVDEPANSKIKESNTESKE-----ENTLEOPKHYED 152
 DB 418 SNSKPTKADFFARLATANGENISDSEETFTESAANSTKMLTFPSSSQOQQOQOP 477
 QY 153 STTSQPSGYTIDKIS-----NODELNLPIINEXENKA-----RPLSTTSAPSIKRYT 202
 DB 478 PKQOQOQOQNHGITSKISAPLNNKKILLRLKNSHISTGAILNNTATISTNPL---- 533
 QY 203 VNOALAEQGSV-----NHLKYTD-----QSTTEGIDSEGYIK 237
 DB 534 -NSVVMQNNNMMSGHMLDELSTIKQEPHQLQOQOPPMQVQSVSYTSDNPSNVIAK 592
 QY 238 AHDAENLIYDTFEDVDKVGSGDMITVIDKNTVPSDLTDSFTTPKINDSGEI-----I 292
 DB 593 SPDRSSSLVLS-----KVSFHLISSTSSNGNTI-----SCPNAVATNSQELEPNNDI 639
 QY 293 AT-----GTYNKNKOITTFEDYVD-----TYENIKAKHLKITS 326
 DB 640 STKSLSLNSTLRHSSANNSNYGDKRPLRTYVSKIFPDSNPGACPLRKYSGPDVNL 699
 QY 327 YIDSKVPNNNTKLDVEYKTLASSYVKTIYEQRPENRFRANLOSMT----- 374
 DB 700 YIEO-----PHTMNSYKCD-----EFYNSRNKKPHGLNFGDNNVIEENNG 746
 QY 375 -TIDPKNHVEQTIYNPLRYSAKETNVNMSGNGDSFTTIDOSTLIKYYKVDNQL 432
 DB 747 DSSVNRPOHTNLQHEF-----PEDNES-----DENDISHSMFYNNHKKNDL 787
 QY 433 PDSNRIDYSEVEDYNDYAQLGNNNDVNIENGINDSFYLIKYSKID--PNKDYT-T 489
 DB 788 ETKPLISIDGEDYV--DDYDR--PNATFNSYGSASNTHELPLGHRMPSRNNNDYDFM 843
 QY 490 IQQVYVMTOTINETY-----GEFTASYDNTIAFTSTSGOGGGL 529
 DB 844 VGNNTGNNQNLNEYTPLRMKRGORHLSTNNNSIMNGSIHMNGNDV 889

RESULT 13
 TANA_XENLA STANDARD; PRT; 1744 AA.
 ID TANA_XENLA
 AC 001550;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Tabinin.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
 CC Xenopodidae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Tadpole head;
 RX MEDLINE=9238961; Pubmed=1524825;
 RA Hemmati-Britvanlou A., Mann R.W., Harland R.M.;
 RT "A protein expressed in the growth cones of embryonic vertebrate
 RT neurons defines a new class of intermediate filament protein."
 RL Neuron 9:417-428(1992).
 CC -1- TISSUE SPECIFICITY: GROWTH CONES OF EMBRYONIC VERTEBRATE NEURONS.
 CC -1- DEVELOPMENTAL STAGE: IS EXPRESSED IN THE NEURULA AND PERSIST
 CC DURING EMBRYOGENESIS IN THE BRAIN, CRANIAL NERVES, AND SPINAL
 CC CORD.
 CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M99387; AAA49966.1; -
 DR PTR: JH0720.
 DR InterPro: IPR001664; IF.
 DR Pfam: PF00038; Filament; 2.
 DR ProSite: PS00226; IF; 1.
 KW Intermediate filament; Coiled coil; Neutrone.
 FT DOMAIN 1 12 HEAD.
 FT DOMAIN 13 314 ROD.
 FT DOMAIN 315 1744 TAIL.
 FT DOMAIN 8 48 COIL 1A.
 FT DOMAIN 49 60 LINKER 1.
 FT DOMAIN 61 156 COIL 1B.
 FT DOMAIN 157 179 LINKER 12.
 FT DOMAIN 180 193 COIL 2A.
 FT DOMAIN 194 199 LINKER 2.
 FT DOMAIN 200 314 COIL 2B.
 SEQID 1744 AA; 199561 MW; 6502EAC9FE6C4E93 CRC64;

Query Match 6.0%; Score 184.5; DB 1; Length 1744;
 Best Local Similarity 19.7%; Pred. No. 0.83; Indels 245; Gaps 42;
 Matches 149; Conservative 122; Mismatches 239;

QY 8 SDESEKNDVNNOSINTDNNQIILKEETNN--VDGIKESDRTSTTVNDEATFL 65
 DB 886 SEQLEKDEINDEQLKQDQIREADTFEVHQVYDFQDQSFEEVQGLNNIKKEVYL 945
 QY 66 QKTPD-----NTHLTEEVR--ESSV-----ESSNSIDT- 95
 DB 946 QNYDEDSQNNDEPQELSCDQEQIKLEENQLSNEGNOFNGNDIEFSQGGYDTD 1005
 QY 96 -----AQPSHTTINREE-----SVQTSQNVDSHVSDFANSKIKESN 133
 DB 1006 EICQETINQVSAQLCESDINQKLSMEDDEEONNPETEDIGLEQSDQENTRSNGT 1065
 QY 134 TESGKEETIEQPNKVKEDSTTSQPSGYTIDKISNO--DELLNLPINE--YENKARPL 189
 DB 1066 KFSQGECCVVERP-----BD--MSDKSEYSGQOEDLDKQVTGFSLNEQANNLDLEKEVYL 1119

```

OY 190 SRTSAOPSI-----KRVTVNOLAEOGSVNNHLIKYTDOSI---REGY----- 229
DB 1120 HHADQDSVNDITDEKLSERITIDNELAT---VDVNSLAANKQVLDLFEYVDVNDV 1176
OY 230 ---DDESEGVAKAH-----DAENLIYDVTFE-----VDDKRVKSGDPTMYDI 266
DB 1177 GMQDDSDSQYQKEDLFDVGNIIKIEIQOTSLNQEICERVDVDEDI--SGEAKNESV 1235
OY 267 DKNVTP-----SDLTDS---FTIPKIDNGSEIIATGYDNKNKOITYTF 308
DB 1236 EKNQVVDVLPKAVYGDQISPLQDEKINLETMETDKNDQGL-----CLEKENE----- 1285
OY 309 TDYVAKYENIK-----AHLKLTSYIDSKVP-----NNNRKL-----DVEKTLASS-- 350
DB 1286 TEYIEVTSQFPQATLSDHAGREL--VDQNSANLOCFENPKRLIAHHIELEYADSDL 1343
OY 351 -----VAKTIVEYORPENNTANIQSMFTNIDTKNHTVEQTIYINPLKYSKETTNYI 404
DB 1344 ESTEQVQETERIPF-KPEDSKMEN-----ENSESESVDSQSEISLN-----SHKSEPEI 1393
OY 405 SGNDEGSGTIIDSTIIVKYKGNQNLPSNRITYD--SEYEDVTN--DDYAQLGNNDV 461
DB 1394 S-----KDYOL--EQTLPDVPLPNELEDFEDLCEQPDVHEHQNNDS 1435
OY 462 NINFGNIDSPYIKYISKYDKPKDYYTIIQOTVMTQITNET--GEFRASVDNTIATST 520
DB 1436 GA-----STFTTV-----DEKEREERESVSDEESNEEFQDVLVDKTSQVEVTT 1483
OY 521 SSGQGGDLPPKTYKIGDYWEDVDKQIQNTND--NEKPLSNV----- 564
DB 1464 LSGLMQ-----EPST-LQD--NESESDSMENAEILNPSNDLYDFMWSQMTETKIITA 1534
OY 565 -----VTLYTPDGTSK---SVRTDEGKYQFD 588
DB 1535 EQVTEQTEVTLQFDDAPNKLFTENLNAKERETYDYE 1569

RESULT 14
Y4A_YEAST STANDARD: PRT: 817 AA.
AC P46949;
AD 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Hypothetical 90.8 kDa protein in RRP41-SNG1 intergenic region.
YGR196C OR G7589.
Saccharomyces cerevisiae (Baker's yeast).
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN NCBI_TaxID=4932;
RP SEQUENCE FROM N.A.
RC STRAIN=S288C;
RX MEDLINE=96076633; PubMed=7502584;
RA Guerrero I., Maita e Silva A., Barreiros T., Arroyo J.,
RA Garcia-Gonzalez M., Garcia-Saez M.I., Rodrigues-Pousada C.,
RA Nombela C.;
"the complete sequence of a 9000 bp fragment of the right arm of
RT Saccharomyces cerevisiae chromosome VII contains four previously
RT unknown open reading frames."
RU Yeast 11:1087-1091(1995).

-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X82775; CAA58019.1;
DR EMBL; Z72981; CAA97222.1;

```

```

DR SGD: S0003428; YGR196C.
KW Hypothetical protein.
SQ SEQUENCE 817 AA; 90797 MW; E52CD659D63BEBB CMC64;

Query Match
Best Local Similarity 21.0%; Pred. No. 0.36;
Matches 134; Conservative 88; Mismatches 237; Indels 180; Gaps 31;

OY 8 SDEEKNVYINN-----NQSINTDNNQI--IKKEETNNVDGIEKRSDD 49
DB 120 SEHSKSDYLSTASLAKSPSEKKSPTHNRFAVNEIDLNLQISREMPPE---IROTSDF 176
OY 50 RTESTTNDE--NEATFLQKTPQDNTHLTEEEYKSSSVSSNSSIDPAQPSHTTINE 107
DB 177 RRDS--DSCDEIQNEAPLCEAVPSSSPVEDEKSHSLGSMDTNADTFE--NTPTRGN 233
OY 108 ESVQSDNVEDSHVSDPANSKIKESN-----TESGKEENTIBQPKVREDSTTSOPGYT 162
DB 234 EHLSSDGVSEKQKDEF---KYSERGIADILPAKKEEN-LQO-----EDQEVESGAL 284
OY 163 NIDEKISNDELNLPLINEYENKARPLSTSAOPSI---KRVTVNOLAEOGSVNNHLIK 219
DB 285 EKKESEKETSIRN-----RN-----STSGQDKYAKPKPVANETKTSIDNGYRMSFMD 333
OY 220 VTDSITFGYDSDSEGVAKAHDAENLIYDTEFVDKVKSGDITVTDIKNTVPSDLTD-- 277
DB 334 YQHSSDSEEDNNEGSSSD-----DNRSSVSQKHADINKQSLDTTDD 381
OY 278 --SFTIPKIDNGSEIIATGYDNKNKOITYFTDYVAKYENIKAHKLTSYIDSKYVN 335
DB 382 ALSTY-ESIKYSTNTEEDNEDNESIE-----DKNE-----DNESIED 419
OY 336 NNTKLDEYKTAALSSVNTTIVEYQRPNE-----RTANLQSMFTNIDTKN 381
DB 420 ENEDTD-SYKFSNRKGSILTSDEEEKGMSDSEGLKAPKSGYFSKMGN--DDKG 477
OY 382 HTVEQTIYINPLKYSKETTNYISGNGDEGSTIIDSTIIVKYKGNQNLPSNRITYD 441
DB 478 DSALQPNQIDTI---EWNLSNSGSELENSDGSDEEHINEDKYLEESSVSDSDVDVSW 533
OY 442 SEYEDVTNDYKQLGNNDVNINFGNIDSP--YII-----RVISKYD-- 481
DB 534 KPDSALNSGFVQ-----DTANKKAPPGIVLDSNGKLVLDLPASAKPRVYSTISM 584
OY 482 -----PNK---DDYTIQOTVMTQITINEYTGFEFRASVDNTIATSTSGQ--QGDL 530
DB 585 ESTWDAFPSKGEDDLLETIRDTKI-----YNNNTIYVNPGLIGNSNLP 629
OY 531 P-----EKTYYKIGDYWEDVDKQIQNTNDNKPPLSNVL 564
DB 630 PLPMDAQEQOLNAGN-----DNSTTDNDSNNTNANDL 660

RESULT 15
RBP2_PLAVB STANDARD: PRT: 1251 AA.
AC Q00799;
AD 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Reticulocyte binding protein 2 (Fragment).
GN RBP2.
OS Plasmodium vivax (strain Belen).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=31273;
RN NCBI_TaxID=31273;
RP SEQUENCE FROM N.A.
RX MEDLINE=92315338; PubMed=1617731;
RA Galinski M.R., Medina C.C., Ingravallo P., Barnwell J.W.;
RT "A reticulocyte-binding protein complex of Plasmodium vivax
RT merozoites."
RU Cell 69:1213-1226(1992).

```

Search completed: July 30, 2002, 10:41:04
Job time: 548 sec

```
CC -1- FUNCTION: INVOLVED IN RETICULOCYTE ADHESION. SPECIFICALLY BINDS TO
CC HUMAN RETICULOCYTE CELLS.
CC -1- SUBCELLULAR LOCATION: Membrane-bound (Probable).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M88098; AAA29744.1; .
KM Malaria; Receptor; Membrane.
FT NON_TER 1251 1251
SQ SEQUENCE 1251 AA; 143741 MW; 54BA51C7404AC572 CRC64;

Query Match 6.0%; Score 184; DB 1; Length 1251;
Best Local Similarity 20.0%; Pred. No. 0.59;
Matches 140; Conservative 115; Mismatches 236; Indels 210; Gaps 37;

UY 18 NNNOSINTDNNQIIRKEE--TNNYDGIERKSRDPR-----STTNV--DENEARFLQK 67
DB 302 SSSNKNVINEINENINRNSQYIKDIEDAKQASTKVELFKHETTSNIFKESSEILGVET 361
OY 68 TPQDNTHLTEEVKSSSVSSSSSIDTPAQPSTHTTINREESVQSDNVEDSHVSD-FAN 126
DB 362 KQOKINKKAEIDIMKE--IERHNSIEIQGVKQFQENLKNLNEPHYDNMADELNDKSTN 418
OY 127 SKIK-ESNTESGKEENTIQPKVKEDSTTSQPSGTYNIDEKISHODELLNPINEYENK 185
DB 419 AKVLEETNLESVK--HNLESETINIKQG-----EKYSK-----AKDIMOK 457
OY 186 ARPLSTTSQPSIKRYTVNQALAEQGSNNHLIKVTDS--ITE-----GYDGS-----E 233
DB 458 IKATSENTAEKTLKY-----KDDOSNYVNYLQITTEKRLNLYTEKNRLNGIDSTITNE 512
OY 234 GVIAKHADENLIVDTF-EVDDKVKSGDTMTVIDIKNTVP-----SDLTDSFTIPKIKD 286
DB 513 GALKESKGN--YEIGFLEKLEIGKRNRLKYDITKKSINSFVGNFSSLFNNFDLNOYDF 569
OY 287 NGEIITATGYONKNQIYTFIDYVDK-ENIK-----AHLKLTSYIDSKY 333
DB 570 NKN---INDYENKKEIYNEFEGLNKISENLNASENTSDYNSAKTLRLAQEKQVNL 625
OY 334 PN-----NNTKLDVE-----YKTAASSVKNKTITVEYQRPNNRPTANLQSMFTNID 378
DB 626 LKKEEENKYLRYKVESFRFPFNKESLQINEMIKKEQLVNEGH-GNVKQVLENIK 684
OY 379 T-----KHHVTEQTIYINPLRYSAKETNVNISGNGDEGSTIIDSTTIK 422
DB 685 ELVDENNLSDILKQATGKNEEI-QKITHSTLKNKAK---TILGHVDTSAKYVG---IK 735
OY 423 V-----YKVGDNQNLPSDSNRIYDSEYEDVTNDYVQALGNNDVNINFGNIDS 470
DB 736 ITPPELALTELLGDAKLKTAQELFEKSKNNVLTENNKK-----NTNELDVH-KNIQD 787
OY 471 PY--IIKVISKYD-----PNKDDYTTIO--QVTMOTITNEYTGFEFT 509
DB 788 AYKVALEIILAHSEIDITKQDSSKLIEMGNQIILKVVLLINQYKKNKISSISKEAIVSVKI 847
OY 510 -----ASYDNTIAF--STSSGOGGDLPEPEKTYKIGDYVWEDVDKGI 550
DB 848 GNVSKKHSELSKITCSGSKSYDNITALEKQTELNLNRSFTQEK-----891
OY 551 QNTNDNEK-----PLSNVLVTLTYPDGTSKSVRTDED 582
DB 892 -NTNSDSKLEKIKTDFESLKNALKTL--EGEVNALKASSD 928
```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 30, 2002, 10:31:01 ; Search time 88.07 Seconds

(without alignments)
1164.823 Million cell updates/sec

Title: US-09-147-405-11

Perfect score: 3087
Sequence: 1 HHHHHHPSDEKNDVNNNN.....SKSVRTDEDKYQPDGQVD 593

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: SPREMBL_19:*
2: SP archaea:*
3: SP bacteria:*
4: SP fungi:*
5: SP human:*
6: SP invertebrate:*
7: SP mammal:*
8: SP mhc:*
9: SP organelle:*
10: SP phage:*
11: SP plant:*
12: SP rodent:*
13: SP virus:*
14: SP vertebrate:*
15: SP unclassified:*
16: SP virus:*
17: SP bacteriophage:*
17: SP archaea:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3015	97.7	1092	2	070022
2	2812.5	91.1	931	2	09K113
3	1162.5	37.7	1171	2	09K113
4	1146.5	37.1	1166	2	086489
5	1139.5	36.9	1141	16	099W46
6	1139.5	36.9	1141	16	099W46
7	576	18.1	1315	2	086488
8	559	18.1	1385	16	099W47
9	506	15.4	1733	2	09K114
10	490	15.9	953	16	099W48
11	460.5	14.9	881	2	093M17
12	456	14.8	947	2	086487
13	452	14.6	933	2	053653
14	452	14.6	935	16	093205
15	452	14.6	989	16	099V14
16	405.5	13.1	877	16	099R07

17	401.5	13.0	913	2	086476	086476 staphylococ
18	376	12.2	940	2	053682	053682 staphylococ
19	345.5	11.2	961	16	099RD3	099RD3 staphylococ
20	326	10.6	1038	16	099RD2	099RD2 staphylococ
21	273.5	8.9	1038	2	09AER7	09AER7 staphylococ
22	261	8.5	1161	2	09X3M7	09X3M7 streptococ
23	239.5	7.8	1227	2	093VY6	093VY6 staphylococ
24	239	7.7	1039	2	P72534	P72534 streptococ
25	234	7.6	604	5	Q26021	Q26021 plasmodium
26	231.5	7.5	2647	5	Q9U4X0	Q9U4X0 plasmodium
27	225	7.3	1463	2	086919	086919 staphylococ
28	222	7.2	1711	5	077322	077322 plasmodium
29	220	7.1	2843	5	Q96315	Q96315 dictyostell
30	218.5	7.1	518	12	Q9DH49	Q9DH49 amastaxa moo
31	218	7.1	1928	5	Q9U0H2	Q9U0H2 plasmodium
32	218	7.1	2081	10	Q9LH98	Q9LH98 arabidops1s
33	217.5	7.0	559	5	Q9U3Y8	Q9U3Y8 plasmodium
34	216.5	7.0	3394	5	Q77384	Q77384 plasmodium
35	215.5	7.0	1395	2	Q9A1S0	Q9A1S0 staphylococ
36	215.5	7.0	1564	2	Q931M3	Q931M3 staphylococ
37	213.5	6.9	2206	5	Q96205	Q96205 plasmodium
38	213	6.9	821	3	Q08581	Q08581 saccharomyc
39	212.5	6.9	1125	16	Q9BPM9	Q9BPM9 mycoplasma
40	212.5	6.9	3130	5	Q9BK46	Q9BK46 plasmodium
41	212	6.9	1946	5	Q97291	Q97291 plasmodium
42	212	6.9	2771	5	Q26216	Q26216 plasmodium
43	211.5	6.9	1344	2	Q49545	Q49545 mycoplasma
44	211.5	6.9	3724	5	Q77320	Q77320 plasmodium
45	211	6.8	691	5	Q9GSD1	Q9GSD1 plasmodium

ALIGNMENTS

RESULT 1
070022 PRELIMINARY; PRT; 1092 AA.

AC 070022;
DT 01-AUG-1998 (TREMblrel. 07, Created)
DT 01-AUG-1998 (TREMblrel. 07, Last sequence update)
DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
DE FIBRINOGEN-BINDING PROTEIN PRECURSOR.
OS Staphylococcus epidermidis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
CC Bacillus/Staphylococcus group; Staphylococcus.
OX NCBI_TaxID=1282;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HB;
RX MEDLINE=98261511; PubMed=9596732;
RA Nilsson M., Frykberg L., Flock J.I., Pel L., Lindberg M., Guss B.;
RT "A fibrinogen-binding protein of Staphylococcus epidermidis.";
RL Infect. Immun. 66:2666-2673(1998).
DR EMBL: Y17116; CAA76638.1; -;
DR InterPro: IPR001899; Gram_pos_anchor.
DR Pfam: PF00746; Gram_pos_anchor: 1.
DR PROSITE: PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.
KW Signal.
FT SIGNAL
FT CHAIN
SQ SEQUENCE 1092 AA; 119292 MW; 6542BC39AAD8B984 CRC64;

Query Match 97.7%; Score 3015; DB 2; Length 1092;
Best Local Similarity 99.7%; Pred. NO. 6.2e-113;
Matches 582; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 8 SDEKKNVNNOSINDDNNQIIKKETNNYDIERKRSERFESTNNVDEENATFLQK 67
DB 75 SDEKKNVNNOSINDDNNQIIKKETNNYDIERKRSERFESTNNVDEENATFLQK 134
OY 68 TPODTHLTREEVKSSSESSNSSIDTAOQPSHTTINRESVQTSADVEDSHVSDFANS 127
|||||

Db 135 TPQDNTHLTEEEVKSSSVSSSSSSIDTAQPSHTTINRESVQTSNDVEDSHVSDFANS 194
Qy 128 KIKESNTESGKEENTIEOPNKKVEDSTTSQPSGYINIDEKISNOBELNLPINEYENKAR 187
Db 193 KIKESNTESGKEENTIEOPNKKVEDSTTSQPSGYINIDEKISNOBELNLPINEYENKAR 254
Qy 188 PLSTTSAOPSIKRVTVNOLAAEOGSNNVHLIKVTDOSTIEGYDSEGVYKADAEENLIYD 247
Db 255 PLSTTSAOPSIKRVTVNOLAAEOGSNNVHLIKVTDOSTIEGYDSEGVYKADAEENLIYD 314
Qy 248 VTFEVDKVKSGDITVTDIDKNTVPSDLTDSFTPIKIDNSGEIATGTVDKKNKQIYTT 307
Db 315 VTFEVDKVKSGDITVTDIDKNTVPSDLTDSFTPIKIDNSGEIATGTVDKKNKQIYTT 374
Qy 308 FTDYVDKXENIKAHKLKLSYIDSKSVNNKTKLVEYKTAALSVKKTITVEYQRPENERT 367
Db 375 FTDYVDKXENIKAHKLKLSYIDSKSVNNKTKLVEYKTAALSVKKTITVEYQRPENERT 434
Qy 368 ANLOSMFTNIDTKNHTVEOTIYINPLRSKAKETNNVINGNDEGSTIIDDSTIIKVKYVG 427
Db 435 ANLOSMFTNIDTKNHTVEOTIYINPLRSKAKETNNVINGNDEGSTIIDDSTIIKVKYVG 494
Qy 428 DNQNLPSNRIYDYSEEDVTNDYAOLGNNNDVNIINGNIDSPIIKIVISKYDPNKDXY 487
Db 493 DNQNLPSNRIYDYSEEDVTNDYAOLGNNNDVNIINGNIDSPIIKIVISKYDPNKDXY 554
Qy 488 TTIOQTVMQTTINEYTEGFRTASDNTIAFSTSSGOGGDLPPPKYTKIGDYWEDVDK 547
Db 555 TTIOQTVMQTTINEYTEGFRTASDNTIAFSTSSGOGGDLPPPKYTKIGDYWEDVDK 614
Qy 548 DGIQNTNDNEKPLSNVLTLYTPDGTSSKSVRTDEGKYQDPGVQ 591
Db 615 DGIQNTNDNEKPLSNVLTLYTPDGTSSKSVRTDEGKYQDPGLK 658

RESULT 2
Q9K113
ID Q9K113 PRELIMINARY; PRT: 931 AA.
AC Q9K113:
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE FIRINGEN-BINDING PROTEIN SDRG.
GN SDRG.
OS Staphylococcus epidermidis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Staphylococcus.
NCBI_TaxID=1282;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN-K28;
RX MEDLINE=20340957; PubMed=10878118;
RA McCrea K.W., Hartford O., Davis S., Ni Eidhin D., Lina G.,
RA Spziale P., Foster T.J., Hook M.;
RT "The serine-aspartate repeat (Sdr) protein family in Staphylococcus
epidermidis".
RL Microbiology 146:1535-1546(2000).
DR EMBL: AF245042; AAF72510.1.
DR InterPro: IPR001899; Gram_pos_anchor.
DR Pfam: PF00746; Gram_pos_anchor; 1.
DR PROSITE: PS00343; GRAM_POS_ANCHORING; UNKNOWN.1.
SQ SEQUENCE 931 AA; 102955 MW; 591E657D97027116 CRC64;

Query Match 91.1%, Score 2812.5; DB 2; Length 931;
Best Local Similarity 93.3%; Pred. No. 6,2e-105;
Matches 545; Conservative 17; Mismatches 21; Indels 1; Gaps 1;

Qy 8 SDEKKNVNNOSINTDNNQIIKKEETNNYDGIKRSERSTSTNNVDENAEFTLOK 67
Db 74 SSNEKNNDVINSQISINTDDNQ-IKKEETNSNDIAENKSDIYQSTNNVDENAEFTLOK 132
Qy 68 TPQDNTHLTEEEVKSSSVSSSSSSIDTAQPSHTTINRESVQTSNDVEDSHVSDFANS 127

Db 133 TPQDNTHLKEEYKPSVSSSSSSSSMDTAQPSHTTINSEASTQTSNDNEENSRVDFANS 192
Qy 128 KIKESNTESGKEENTIEOPNKKVEDSTTSQPSGYINIDEKISNOBELNLPINEYENKAR 187
Db 193 KIKESNTESGKEENTIEOPNKKVEDSTTSQPSGYINIDEKISNOBELNLPINEYENKAR 252
Qy 188 PLSTTSAOPSIKRVTVNOLAAEOGSNNVHLIKVTDOSTIEGYDSEGVYKADAEENLIYD 247
Db 253 PLSTTSAOPSIKRVTVNOLAAEOGSNNVHLIKVTDOSTIEGYDSEGVYKADAEENLIYD 312
Qy 248 VTFEVDKVKSGDITVTDIDKNTVPSDLTDSFTPIKIDNSGEIATGTVDKKNKQIYTT 307
Db 313 VTFEVDKVKSGDITVTDIDKNTVPSDLTDSFTPIKIDNSGEIATGTVDKKNKQIYTT 372
Qy 308 FTDYVDKXENIKAHKLKLSYIDSKSVNNKTKLVEYKTAALSVKKTITVEYQRPENERT 367
Db 373 FTDYVDKXENIKAHKLKLSYIDSKSVNNKTKLVEYKTAALSVKKTITVEYQRPENERT 432
Qy 368 ANLOSMFTNIDTKNHTVEOTIYINPLRSKAKETNNVINGNDEGSTIIDDSTIIKVKYVG 427
Db 433 ANLOSMFTNIDTKNHTVEOTIYINPLRSKAKETNNVINGNDEGSTIIDDSTIIKVKYVG 492
Qy 428 DNQNLPSNRIYDYSEEDVTNDYAOLGNNNDVNIINGNIDSPIIKIVISKYDPNKDXY 487
Db 493 DNQNLPSNRIYDYSEEDVTNDYAOLGNNNDVNIINGNIDSPIIKIVISKYDPNKDXY 552
Qy 488 TTIOQTVMQTTINEYTEGFRTASDNTIAFSTSSGOGGDLPPPKYTKIGDYWEDVDK 547
Db 553 TTIOQTVMQTTINEYTEGFRTASDNTIAFSTSSGOGGDLPPPKYTKIGDYWEDVDK 612
Qy 548 DGIQNTNDNEKPLSNVLTLYTPDGTSSKSVRTDEGKYQDPGVQ 591
Db 613 DGIQNTNDNEKPLSNVLTLYTPDGTSSKSVRTDEGKYQDPGLK 656

RESULT 3
Q9KMX6
ID Q9KMX6 PRELIMINARY; PRT: 1171 AA.
AC Q9KMX6:
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE BONE SIALOPROTEIN-BINDING PROTEIN.
GN BBP.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Staphylococcus.
NCBI_TaxID=1280;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN-024;
RX MEDLINE=20115096; PubMed=10642520;
RA Tung H.S., Guss B., Hellman U., Persson L., Rubin K., Ryden C.;
RT "A bone sialoprotein-binding protein from Staphylococcus aureus: a
member of the staphylococcal Sdr family".
RL Biochem. J. 345:611-619(2000).
DR EMBL: Y18653; CAB75732.1.
DR InterPro: IPR001899; Gram_pos_anchor.
DR PROSITE: PS00343; GRAM_POS_ANCHORING; UNKNOWN.1.
SQ SEQUENCE 1171 AA; 127123 MW; C5BC812F9DA5A884 CRC64;

Query Match 37.7%, Score 1162.5; DB 2; Length 1171;
Best Local Similarity 42.8%; Pred. No. 4,4e-39;
Matches 271; Conservative 101; Mismatches 182; Indels 79; Gaps 20;

Qy 8 SDEKKNVNNOSINTDNNQIIKKEETNNYDGIKRSERSTSTNNVDENAEFTLOK 58
Db 57 STENAKODEASASDNKEVSESTENSTQKNDLTPNKRKTVDSHQEKKEAPTSSSTQOQ 116
Qy 59 ENAEFTLOKTPQDNTHLTEEEVKES---SVSSSSSSIDTAQPSHTTINRESVQTSND 115

```

Db 117 ONNATSTETKPN--IEKENVKPSTDKTAFEDTSVILEKKAPANT--NNDVTRKPS-- 170
Qy 116 VEDSHVSPANSKIKESNT---ESGKEENTIEQPNKVKED---STTSQSGYTN--DEK 167
Db 171 -----TSEIQTPTPTQESTNINSOPQPPSKYVDNVDJTTNKEPVNSKEBL 220
Qy 168 ISNOELNLPINE--YENKARPLSTSAQPSIKRV-----TVNOLAEOGSNNHLIK 219
Db 221 KNBPKELELRNDSNTSTRTPVATAPTVAAPKRVNAKIRPAVAQPAVANNNDLIT 280
Qy 220 VTPDSITEGYDSEGVAKAHDAENLIYDVFEDDKVKSQDGMTVVDIDKNTPVSDLTDSF 279
Db 281 VTKQITEGIND-DGVIQAHQSEHITYSDFRIDNAKAGDMTKYDKETIPSDITDDE 339
Qy 280 TTPKIKDSGEIATGTYDNKNKQITTYFTDYVDYKENIKAKHLKLSYIDSKVPPNNMTK 339
Db 340 TPVDITDSGGEYIAKGTFDLNTKTYKFTDYVDREYVNAKLEINSYDKKEVP--NETN 398
Qy 340 LDVEKKTALSSVNTITVEYORPNENRTANLQSMETNIDTKNHTVEQTIYINPLRYSAKE 399
Db 399 LNLTPATADKETSKNVKVEYQKPIYKDESNIQSIFSLDTTQHEVEQTIYVPLKLNKN 458
Qy 400 TNVNI-----SGNGD-----EGSTIIDSTIIIVKYKVDNQNLDPENRIYDSEYEDVTND 450
Db 459 TNVTIKSGGVADNGDYTGDSSTIIDSNTKIKVYKVASGQQLPQSNKITYDSQYEDVINS 518
Qy 451 DYALQGN--NDVINFGNIDSPYIKYISKYDNKDDYTTIOQTVMOTITNEYTGEF 507
Db 519 --VTKNKYGTGMANINCGDIDSAIYKVKVSKYTPGAEEDDLAVOOGVRMTT-----TNKY 571
Qy 508 RTASVDN-----TAFSTSGOGGQDLPEPKTYKIGDYVWEDVDKQIONTNDKX 558
Db 572 NISSAGTITLTYQLTLTVVTVSVK-----PEEKLYKIGDYVWEDVDKVGQGTDSK 626
Qy 559 PLSNVLVTLTYPDGTSKSYRTDEDKYQFDGYO 591
Db 627 PMANVLVTLTYPDGTTKSVRTDANGHYEFGGLK 659

RESULT 4
086489 PRELIMINARY: PRT: 1166 AA.
AC 086489:
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE SDR E PROTEIN.
GN SDR E.
OS Staphylococcus aureus.
Bacteria; Firmicutes; Bacillus/Clostridium group;
Bacillus/Staphylococcus group; Staphylococcus.
NCBI_TaxID=1280;
RA [1]
RN SEQUENCE FROM N.A.
RC STRAIN-NEWMAN:
RX MEDLINE=99098700; PubMed=9884231;
RA Josephson E., McCrea K., Nl Eldhlm D., O'Connell D., Cox J., Hook M.,
RA Foster T.D.;
RT "Three new members of the serine-aspartate repeat protein multigene
RT family of Staphylococcus aureus.";
RL Microbiology 144:3387-3395(1998).
DR EMBL, AJ005647; CAA06652.1;
DR InterPro; IPR001899; Gram_Pos_anchor.
DR Pfam; PF00746; Gram_Pos_anchor.1.
DR ProSITE; PS00343; GRAM_POS_ANCHORING; UNKNOWN.1.
DR PROSITE; PS00343; GRAM_POS_ANCHORING; UNKNOWN.1.
SEQUENCE 1166 AA; 126548 MW; 750A7B0135287D4A CRC64;

Query Match 37.1%; Score 1146.5; DB 2; Length 1166;
Best Local Similarity 42.5%; Pred. No. 1.9e-38;
Matches 263; Conservative 110; Mismatches 191; Indels 55; Gaps 18;
Oy 8 SSDEEKNDVYINNQSINTDNNQI--IKKEETNNYDGIKRSRSDRPRESTTNDVENDATFL 65

```

```

Db 66 ATTSNKEVESETEENNSTTENNSTNPIKKE--TNTDSQPEAKKESTSSSTQKQNNVTAT 123
Qy 66 OKTPDNDNLTREEYKES--SSVESSNSIDTAAQPSHTTINREESVQTSQDNVEDSHV 122
Db 124 TETKPN--IEKENVKPSTDKTAFEDTSVILEKKAPANT--NNDVTRKPS-----TS 172
Qy 123 DEANSKIKESNT---ESGKEENTIEQPNKVKEDSTTSQSGYTN--DEK 167
Db 173 EPSTSEIQTPTPTQESTNINSOPQPPSKYVDNVDJTTNKEPVNSKEBLKMP 229
Qy 179 -----INEYENKARPLSTSAQPSIKRV-----TVNOLAEOGSNNHLIKYVDQ 223
Db 220 EKLKELVRNDSNTSTRTPVATAPTVAAPKRVNAKIRPAVAQPAVANNNDLITKYKQ 289
Qy 224 SITEGYDSEGVAKAHDAENLIYDVFEDDKVKSQDGMTVVDIDKNTPVSDLTDSFTIPK 283
Db 290 TIKVG-DGKDNVAAHDKGDIYEFTIDNKVKKGDMTINVDKNVIPSDLTDRNDPID 348
Qy 284 IKDSNGEIIATGTYDNKNKQITTYFTDYVDYKENIKAKHLKLSYIDSKVPPNNMTK 343
Db 349 ITDPGGEVIAKGTFDKATKQITTYFTDYVDYKEDIKSRLLTYSYIDKKTVP--NETSLMLT 407
Qy 344 YKTALSSVNTITVEYORPNENRTANLQSMETNIDTKNHTVEQTIYINPLRYSAKETVNN 403
Db 408 FATACKETSQNTVYQDPMVHGDNSIQTFTKDEDKOTIEQIYVPIKKSATNTKYD 467
Qy 404 ISGNG-----DEGSTIIDSTIIIVKYKVDNQNLDPENRIYDSEYEDVTND-DYA 453
Db 468 IAGSOVDYGNKIKLNGSGSTIIDQNEIKYKVNQDQLPQSNRIYDFQYEDVTSQFQPNK 527
Qy 454 QLGNNNDVINFGNIDSPYIKYISKYDNKDDYTTIOQTVMOTITNEYTGEFTAYD 513
Db 528 KSESNVAVATLDEGIDNSAIIKRVSKYTPSDGELIDAGTSMRT--DKY-GIYVAGS 585
Qy 514 NTIAFSTSGOGGQDL-PEPKTYKIGDYVWEDVDKQIONTNDKXPLSNVLVTLTYPDG 572
Db 586 NPIVNSNDTGGGQDYKPEKIKYKIGDYVWEDVDKQIONTNDKXPLSNVLVTLTYPDG 645
Qy 573 TSKSVRTDEDKYQFDGYO 591
Db 646 TTKSVRTDANGHYEFGGLK 664

RESULT 5
099W46 PRELIMINARY: PRT: 1141 AA.
AC 099W46:
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE SER-ASP RICH FIBRINOGEN-BINDING, BONE STALOPROTEIN-BINDING
DE PROTEIN.
GN SDR E OR SA0521.
OS Staphylococcus aureus (strain N315).
Bacteria; Firmicutes; Bacillus/Clostridium group;
Bacillus/Staphylococcus group; Staphylococcus.
NCBI_TaxID=158879;
RA [1]
RN SEQUENCE FROM N.A.
RC MEDLINE=23311952; PubMed=11418146;
RX Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.T., Nagai Y., Ito T., Kanamori M.,
RA Matsumaru H., Matuyama A., Murakami H., Hoshoyama A., Mizutani-U Y.,
RA Takahashi N.K., Sawano T., Inoue R.I., Kaito C., Sekimizu K.,
RA Hiraoka H., Kuhara S., Goto S., Yabuzaki J., Kanehisa M.,
RA Yamashita A., Oshima K., Furuya K., Yoshino C., Shibata T., Hattori M.,
RA Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus.";
RL Lancet 357:1225-1240(2001).
DR EMBL, AP003131; BAB41752.1;
DR InterPro; IPR001899; Gram_Pos_anchor.

```

DR Pfam: PF00746; Gram_pos_anchor.1.
 DR PROSITE: PS00343; GRAM_POS_ANCHORING. UNKNOWN.1.
 KW Complete proteome.
 SQ SEQUENCE 1141 AA; 124026 MW; 445419D8BC5A4F8 CRC64;

Query Match 36.9%; Score 1139.5; DB 16; Length 1141;
 Best Local Similarity 43.1%; Pred. No. 3.6e-38;
 Matches 266; Conservative 99; Mismatches 193; Indels 59; Gaps 19;

QY 8 SDEEKNVNNNSINFDNNQIIKKEETNNYDGIKRSRDEESTTNVDENETFLQK 67
 DB 69 SDNKEVSETENNSTENDSTNPI--KKEFTN-DSOPEAKEESTTSSYQOQOONNTATTE 125
 QY 68 TPQDTHLLEEVKES---SSVESSNSIDTAQOPSHTTINREESVOTSDNVEDSHVDF 124
 DB 126 TKPQN--TEKENVKPSTOKTATEDSVILEEKKAPNT--NNDVYTKFS----- 170
 QY 125 ANSKIKESNT---ESGKEENTIEOPNKYKEDSTTSPSGYTNIDEKIS--NODELINLP-- 178
 DB 171 -TSEIOTKPTTPOESTNINENSOQOPTPSKVD---NQVTDATNPKPEPVANSKEELKNNEPEK 226
 QY 179 -----INEXENKARPIST--TSAQP-----SIRKVTYNOLAEGGSVNNHLIKYTDOSI 225
 DB 227 LKELYRNNDNDRSTKPAATAFTSVAPKRLNKKMFAPAQAAPAAVANSNNVNDLITVTKOTI 286
 QY 226 TEGYDSDSGVIAKHADEMLIYDVFPEVDKYSGDTMTVDIDKNTVPSDLTDSFTIPKIK 285
 DB 287 KVG-DGKNVAAAHDKGKIEVDTEFTIDNKKYKGGDTMTINDKNYIPSDLDKNDPIDIT 345
 QY 286 DNSGIIATGTIDNKKKQITTYTIDYVDKYENIKAKHLTSTYIDSKVPPNNNTKLDVEYK 345
 DB 346 DPGSEVIKAGTFDKATKQITTYFTDYVDKYEDIKARLTIYSYDKOAVF--NETSLNTFA 404
 QY 346 TALSSVNNKTIYEVORPENRNTANLOSMEFTNIDTNHVEQITTYINPLRYSAKETNNVIS 405
 DB 405 TAGKETSQNVSVYDQDPVHAGDSNIOSTFTKLDENKQITIEQIYVNLPKTATNTKRVIA 464
 QY 406 GNG-----DEGSTIIDSTIIKYYKVGDNQNLPSNRITYDSEYEDVTND--DYAQL 455
 DB 465 GSQVVDYGNIKLNGSTIITDQNTTEIKYKVNPNQOLPOSNRITYDSQYEDVTSQDNKKS 524
 QY 456 GNNNDVNINFGNIDSPYIIKVISKYDPNKDDYTTIQQVYTMQTTINETGEERTASYDNT 515
 DB 525 FSNMVAATLDFGDISAYIIKIVSKYTPISDGLDIAQGTSMRT--DKY-GYNYVAGYSNF 582
 QY 516 IAFSTSSGQGGDL--PPEKTYKIGDYVWEDVDKGIQNTNDEKPLSNVLTLYTPPDGTS 574
 DB 583 IYTSNDTGGDGTVPPEKLYKIGDYVWEDVDKDVQGTDSKEXEMANVLTLYTPPDGTT 642
 QY 575 KSVRTEDGKYQFDGVQ 591
 DB 643 KSVRTDANGHYEFGGLK 659
 RESULT 6
 ID Q932F7 PRELIMINARY; PRT; 1141 AA.
 AC Q932F7;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE SER-ASP RICH FIBRINOGEN-BINDING, BONE STALOPOTEIN-BINDING
 DE PROTEIN.
 DE GN SDRE OR SAV0563.
 OS Staphylococcus aureus (strain Mu50).
 OC Bacteria; Firmicutes; Bacilli/Clostridium group;
 OC Bacillus/Staphylococcus group; Staphylococcus.
 OX NCBI_TaxID=15878;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21311952; PubMed=11418146;
 RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,

RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Iian J.-Q., Ito T.,
 RA Kanamori M., Matsunaru H., Maruyama A., Murakami H., Hosoyama A.,
 RA Mizutani-Ui Y., Takahashi T., Sawano T., Inoue R.-I., Kaito C.,
 RA Sekimizu K., Hiraoka H., Kubara S., Goto S., Yabuzaki J.,
 RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
 RA Hattori M., Ogasawara N., Hayashi K., Hiratazu K.,
 RT "Whole genome sequencing of methicillin-resistant Staphylococcus
 RT aureus.";
 RL Lancet 357:1225-1240(2001).
 DR EMBL: AP003359; BABS6725.1; .
 KW Complete proteome.
 SQ SEQUENCE 1141 AA; 124038 MW; E679FC2991846D9 CRC64;

Query Match 36.9%; Score 1139.5; DB 16; Length 1141;
 Best Local Similarity 43.1%; Pred. No. 3.6e-38;
 Matches 266; Conservative 99; Mismatches 193; Indels 59; Gaps 19;

QY 8 SDEEKNVNNNSINFDNNQIIKKEETNNYDGIKRSRDEESTTNVDENETFLQK 67
 DB 69 SDNKEVSETENNSTENDSTNPI--KKEFTN-DSOPEAKEESTTSSYQOQOONNTATTE 125
 QY 68 TPQDTHLLEEVKES---SSVESSNSIDTAQOPSHTTINREESVOTSDNVEDSHVDF 124
 DB 126 TKPQN--TEKENVKPSTOKTATEDSVILEEKKAPNT--NNDVYTKFS----- 170
 QY 125 ANSKIKESNT---ESGKEENTIEOPNKYKEDSTTSPSGYTNIDEKIS--NODELINLP-- 178
 DB 171 -TSEIOTKPTTPOESTNINENSOQOPTPSKVD---NQVTDATNPKPEPVANSKEELKNNEPEK 226
 QY 179 -----INEXENKARPIST--TSAQP-----SIRKVTYNOLAEGGSVNNHLIKYTDOSI 225
 DB 227 LKELYRNNDNDRSTKPAATAFTSVAPKRLNKKMFAPAQAAPAAVANSNNVNDLITVTKOTI 286
 QY 226 TEGYDSDSGVIAKHADEMLIYDVFPEVDKYSGDTMTVDIDKNTVPSDLTDSFTIPKIK 285
 DB 287 KVG-DGKNVAAAHDKGKIEVDTEFTIDNKKYKGGDTMTINDKNYIPSDLDKNDPIDIT 345
 QY 286 DNSGIIATGTIDNKKKQITTYTIDYVDKYENIKAKHLTSTYIDSKVPPNNNTKLDVEYK 345
 DB 346 DPGSEVIKAGTFDKATKQITTYFTDYVDKYEDIKARLTIYSYDKOAVF--NETSLNTFA 404
 QY 346 TALSSVNNKTIYEVORPENRNTANLOSMEFTNIDTNHVEQITTYINPLRYSAKETNNVIS 405
 DB 405 TAGKETSQNVSVYDQDPVHAGDSNIOSTFTKLDENKQITIEQIYVNLPKTATNTKRVIA 464
 QY 406 GNG-----DEGSTIIDSTIIKYYKVGDNQNLPSNRITYDSEYEDVTND--DYAQL 455
 DB 465 GSQVVDYGNIKLNGSTIITDQNTTEIKYKVNPNQOLPOSNRITYDSQYEDVTSQDNKKS 524
 QY 456 GNNNDVNINFGNIDSPYIIKVISKYDPNKDDYTTIQQVYTMQTTINETGEERTASYDNT 515
 DB 525 FSNMVAATLDFGDISAYIIKIVSKYTPISDGLDIAQGTSMRT--DKY-GYNYVAGYSNF 582
 QY 516 IAFSTSSGQGGDL--PPEKTYKIGDYVWEDVDKGIQNTNDEKPLSNVLTLYTPPDGTS 574
 DB 583 IYTSNDTGGDGTVPPEKLYKIGDYVWEDVDKDVQGTDSKEXEMANVLTLYTPPDGTT 642
 QY 575 KSVRTEDGKYQFDGVQ 591
 DB 643 KSVRTDANGHYEFGGLK 659
 RESULT 7
 ID O86488 PRELIMINARY; PRT; 1315 AA.
 AC O86488;
 DT 01-NOV-1998 (TREMBLrel. 08, Created)
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE SDRD PROTEIN.
 DE GN SDRD.
 OS Staphylococcus aureus.

Db	553	NALGFTNNQSGAG-----GQFYKKGNTVWEDTNNKNGVQL--GEKGGNTVTVV-PDNTT	603
Qy	574	SKSVR---TDDEGKY	585
		:	
Db	606	NTKVGAVTKEDGGSY	620
RESULT	9		
Q9KII4			
ID	Q9KII4	PRELIMINARY;	PRT; 1723 AA.
AC	Q9KII4		
DT	01-OCT-2000 (TREMBLrel. 15, Created)		
DT	01-OCT-2000 (TREMBLrel. 15, Last sequence update)		
DT	01-JUN-2001 (TREMBLrel. 17, Last annotation update)		
DE	PUTATIVE CELL-SURFACE ADHESIN SDRF.		
GN	SDRF.		
OS	Staphylococcus epidermidis.		
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillus/Staphylococcus group; Staphylococcus.		
	NCBI_taxid=1282;		

.n
 RP SEQUENCE FROM N.A.
 RC [1]
 RX STRAIN=9491.
 RA MEDLINE=20340957; PubMed=10878118;
 RA MCCRae K.W., Hartford O., Davis S., Ni Eidhin D., Lina G.,
 RA Speziale P., Foster T.J., Hook M.;
 RT "The serine-aspartate repeat (sdr) protein family in *Staphylococcus*
 RT *epidermidis*.";
 RL Microbiology 146:1535-1546(2000)..
 DR EMBL; AF245041; AAF72509.1; -
 DR Interpro: IPR001899; Gram_pos_anchor.
 DR PROSITE; PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.
 SQ SEQUENCE 1733 AA; 184720 MW; D8D62EAL692FDE8 CRC64;

Query Match 16.4%; Score 506; DB 2; Length 1733;
Best Local Similarity 27.4%; Pred. No. 9.6e-13;
Matches 174; Conservative 101; Mismatches 276; Indels 84; Gaps 27

[illegible]

D6 592 VKDPTSGAPKPVADGTRVDFINFARSMANCKKLYVTQAVNPRTGICGNVYT--EWLFTDGT 649

Q7 500 INETTGERTASDYNTIAFTSSSSGQGGDLPETKTYIGVWEDVDXDDQNTNDKEP 555

D6 650 TN--TNDYRQTKSTTYIILNGSSTAGQDN--TISGDIYIWLDDKNNGVO--DDEK 702

Q7 560 LSNVLVTLYTPDGTSK--VRIDEDGKGYQPDGVO 591

D6 703 LAGVYVTL--KDSNNRELQRYTTTQSGHYQPDNLQ 735

RESULT	ID	PRELIMINARY;	PRT;	953 AA.
Q99M48	Q99M48			
AC	Q99M48;			
DT	01-JUN-2001 (TREMblrel, 17, Created)			
DT	01-JUN-2001 (TREMblrel, 17, Last sequence update)			
DT	01-DEC-2001 (TREMblrel, 19, Last annotation update)			
DE	SER-ASP RICH FIBRINOGEN-BINDING, BONE STALOPROTEIN-BINDING			
DE	PROTEIN.			
GN	SDRC OR SA0519 OR SAV0561.			
OS	Staphylococcus aureus (strain N315), and			
OS	Staphylococcus aureus (strain M50).			
CC	Bacteria; Firmicutes; Bacillus/Clostridium group;			
CC	Bacillus/Staphylococcus group; Staphylococcus.			
XX	NCBI_TaxID=158879, 158878;			

NC SEQUENCE FROM NCBI.
 RE SPECIES: S. aureus (strain N315), and S. aureus (strain Mu50);
 RX MEDLINE=2131195; PubMed=11418146;
 RA Kuroda M., Ohta T., Uchiyama E., Baba T., Yuzawa H., Kobayashi I.,
 RA Cui L., Oguchi A., Aoki K.-i., Nagai Y., Iian J.-Q., Ito Y.,
 RA Karamori M., Matsunaru H., Maruyama A., Murakami H., Hotoyama A.,
 RA Matsutani U. Y., Takahashi N. K., Sawano T., Inoue R.-i., Kaito C.,
 RA Sekimizu K., Hirakawa H., Kohara S., Goto S., Yabuzaki J.,
 RA Kanehisa M., Yamashita A., Oshima K., Futuya K., Yoshino C., Shiba T.,
 RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
 RT "Whole genome sequencing of methicillin-resistant *Staphylococcus*
 RL aureus".
 RL Lancet 357:1225-1240(2001).
 DR EMBL: AP003131; BAB41750.1; -;
 DR EMBL: AP003359; BAB56723.1; -;
 DR InterPro: IPR001899: Gram_pos_anchor.
 DR Pfam: PF00746; Gram_pos_anchor.1.
 DR PROSITE: PS00343; GRAM_POS_ANCHORING; UNKNOWN.1.
 KW Complete proteome.
 KW SEQUENCE 953 AA; 103292 MW; 729A7169A074A1E5 CRC64;

Query Match	15.98;	Score 490;	DB 16;	length 953;
Best Local Similarity	28.28;	Pred. No. 2.3e-12;		
Matches 162;	Conservative 83;	Mismatches 236;	Indels 94;	Gaps 22

Qy	42	GIKRSDBRTSTTNVNDENATFLQKTPQDNDTHLIEEYKSSSVESNSIDPAOPSH	101
Db	45	GHEAAAHATHGELQSNENET-----ASEKK-----TTEKV--DSQQLKDNTOFATAADPKV	96
Qy	102	TTINREESVQTSDNVEDSHVSDPANSKIKESNTESGKEENTIEOPNKVKEDSTTSOPSGY	161
Db	97	T-----MDSATVTKETSSNMGS-----PQANATASQSTTQRSNV	129
Qy	162	TNIDEK---ISNODELNLPIINEIEN-KARLSSTTSAPSTIKRTYVQOLA-EDGSVYNH	216
Db	130	TTNDSKSTTYNSNETKSNL--TQAKNVSTPTKTTTLIQRALNRNAVTVTAAPQGGTVAND	187
Qy	217	LIKVTDOSI-----TEGYDSESEVIRKAHDAENLIYDTEFEVDKVKSGDTMT	263
Db	188	KVHFINDIALDKGHVKNKTGTEFMAATSSDYLR-----LKANTTIDSVAKEGJFT	239
Qy	264	VDIDKNVPSDLDSFTPKIKDNGSELIATGTDNKNKQITTYTFYDVEYKIEINAKHLK	322
Db	240	EKYGYEFPBGVRLPSPQONLYNNAOGNIANGIYDSKNTNTTYFTTNVVOQYTVSGSFE	299

[illegible][illegible]

```

Db 300 QVAFARAKATTTDKTAYKKEVTLGNDTYSIELIVD---GNKRAQPLISTSTNYINNEDLS 356
Oy 384 VEGTITIN-PLARYSAKETNV-NISGNGDESGTIIIDSTIIKYKVDONNL-----PDSN 436
Db 357 RNMATAVNOPKNTYTKOTETYNLT-----GYKRNPAKNEKITVEVDQNPVDSFPDPS 411
Oy 437 RIYDSEYEDVTDNDYQAOLGNNDVNIINGNIDS--PYIIKYV-----SKYDPNKDYYT 489
Db 412 KIKDVTDQFVY---YSNDKRTATVDLMKGQSTSSNKQYIIIOVAVPDSNDSTGDKIDY-- 466
Oy 490 IQQVMTQTTINTEYGEFRTASTDNTIATSTSGGCGGLDPEPKTKIGDYVMEVDKDG 549
Db 467 -----TLDDDKTKYSW---SNSYSNVNGSSTANGD-----QKRYNLGDYVWEDTNKDG 511
Oy 550 IONTNNEKPLSNVLTLYTPDSTS--KSVPTDEDGKYRPGDV 590
Db 512 KODA--NEKIGKVYILKDSNGEKELDRITTDENGKTOFTGL 551

```

```

RESULT 13
ID 053653 PRELIMINARY: PRT: 933 AA.
AC 053653;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE CLUMPING FACTOR.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Staphylococcus.
OX NCBI_Taxid=1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-NEUMAN;
RX MEDLINE=94224142; PubMed=8170386;
RT Medevitt D., Francois P., Vaudaux P., Foster T.J.;
RT "Molecular characterization of the clumping factor (fibrinogen receptor.
RT of Staphylococcus aureus."
RL MOL. Microbiol. 11:237-248(1994).
DR EMBL: Z18852; CAA7304.1;
DR InterPro: IPR000515; BPD.transp.
DR InterPro: IPR001899; Gram_pos_anchor.
DR PROSITE: PS00402; BPD_TRANSP_INN_MEMBER; UNKNOWN.1.
DR PROSITE: PS00343; GRAM_POS_ANCHORING; UNKNOWN.1.
SQ SEQUENCE 933 AA; 97058 MW; EB51A6DE2FE759F4 CRC64;

```

```

Query Match 14.6%; Score 452; DB 2; Length 933;
Best Local Similarity 26.6%; Pred. No. 7.3e-11;
Matches 154; Conservative 105; Mismatches 234; Indels 86; Gaps 26;

Oy 8 SDEEKNDVYNNOSINTDNNQIIEKEETNNNDGIEKRESDRESTYNDENATFLQK 67
Db 51 SNEKSNDSVSAAPKTDNTN-----SDTKSSNTN---NGETISVAQ 91
Oy 68 TPQDNTHLTREEVKESSVSNSIDTAQOPSHTTINRESVQTSNVEDSHVSDPANS 127
Db 92 NP-----AQOETQSSSTNATTEETPVGTGATTTTNOANTPATQOS--SNTNAEELVN- 143
Oy 128 KIKESNTSGKEENTIEQPNKVKEDSTTSOPSGYTNIDEXISNDELIN-LPINEYENKA 186
Db 144 --QTSNETTNDNTV-----SSVNSPQNSTNAENSTGCDTSTEAATPSN---NES 189
Oy 187 RPLSTSAOSIKRVTVN-----QLAAEGSNVNHILKYVDO--SITEGYDSEGV 235
Db 190 APOSTDSNNDVYNOAVNTSAPRRAPSLAAVAADAPAGTDITNOJLNTVGL--DSGTT 248
Oy 236 IKADHAENLIYDTFEVDKVKSGDTMTVIDKNTVPSDLTSPFIKIDNSGEIATG 295
Db 249 VYRQAGVYKNTGFSVPNSAVKDDTFRITVPRKELNNGVSTAKVPPIMAGD-QVLANG 307
Oy 296 TYDNKKKQITTTFTDYVDKENTIKAKHLITSYIDKSKVPPNNNTKL-DVEIKTALSS--VN 352

```

```

Db 308 VIDS--DGNVIYFTDVTNKKKDVKATLMPAYID-----PENVKKGTGNTLATIGSTTAN 362
Oy 353 KRTVEYQRPENRRTANLQSMETNIDTKNHTVEQTIYINPL--RYSAKETNVNISGNGDE 410
Db 363 KTYLVDEYKGFYKFNLSIGTIDQIDKTNNTYQTIYVPSGQNTAIPVLGTLKRNFTOS 422
Oy 411 GSTIIDSSTIIKYKVGNDQNLDPDSNRITYDSEYEDVTNDYQAOLGNNDVNIING- 466
Db 423 NALIDQONTSIKYKVDNNAADLSESYFV--NPENFEDVTSVNTTFPNNQYKVEFNTPD 481
Oy 467 NIDSPYIIKYSKIDPN-KDQYTIQOYTMQTTINTEYGE--FRASVONTIATFSTSG 523
Db 482 QITTPYIVVNGHIDNSKGD-----LALRSTLGYNSNIITMRSMWDNEVAFNNGSG 534
Oy 524 OGOG-DLP--PEKTKIGDY--VMEVDKDG-GLONTNDN 556
Db 535 SGDGIKPPVYEQPDEPELEIPEPDSDDPSGSDSDS 573

```

```

RESULT 14
ID 0932C5 PRELIMINARY: PRT: 935 AA.
AC 0932C5;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE FIBRINOGEN-BINDING PROTEIN.
GN FNB OR SAV0811.
OS Staphylococcus aureus (strain Mu50).
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Staphylococcus.
OX NCBI_Taxid=158878;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=2111952; PubMed=11418146;
RT Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RT Cui L., Oguchi A., Aoki K.-I., Nagai Y., Iian J.-Q., Ito T.,
RT Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RT Mizutani-Uji Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RT Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RT Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RT Katoori M., Ogasawara N., Hayashi H., Hiramatsu K.,
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus."
RL Lancet 357:1225-1240(2001).
DR EMBL: AP003360; BAB56973.1;
KW Complete proteome.
SQ SEQUENCE 935 AA; 96950 MW; DC5A2D92CE3BA91C CRC64;

```

```

Query Match 14.6%; Score 452; DB 16; Length 935;
Best Local Similarity 25.9%; Pred. No. 7.3e-11;
Matches 152; Conservative 109; Mismatches 224; Indels 102; Gaps 27;

Oy 8 SDEEKNDVYNNOSINTDNNQIIEKEETNNNDGIEKRESDRESTYNDENATFLQK 67
Db 51 SNEKSNDSVSAAPKTDNTN-----SDTKSSNTN---NGETISVAQ 91
Oy 68 TPQDNTHLTREEVKESSVSNSIDTAQOPSHTTINRESVQTSNVEDSHVSDPANS 127
Db 92 NP-----AQOETQSSSTNATTEETPVGTGATTTTNOANTPATQOS--SNTNAEELVN- 143
Oy 128 KIKESNTSGKEENTIEQPNKVKEDSTTSOPSGYTNIDEXISNDELIN-LPINEYENKA 186
Db 144 --QTSNETTNDNTV-----SSVNSPQNSTNAENSTGCDTSTEAATPSN---NES 189
Oy 187 RPLSTSA-----QPSIKRVTVNQLA-----EGGSNVNHILKYVDOSTTEGYD 231
Db 190 APOSTDSNNDVYNOAVNPSPPRRAPSLAAVAADAPAGTDITN--QIDVTKYT---ID 244
Oy 232 SEGVYKADHAENLIYDTFEVDKVKSGDTMTVIDKNTVPSDLTSPFIKIDNSGEI 291

```

```

Db 245 SGTIVPHQAGYKLVNGFSPNSAVKGDFTKIVPEKELNLTGVTSTAKVPPIMAGD-QV 303
QY 292 IATGYDNKKKOITFTDQVVDKXENIKAKLKLTSYIDSKVPNNKTL-DVEYKTL-- 348
Db 304 LANGVIDS-DGNVIYFTDYDNKENVTANTMPAYID----PENVTKGAVTLTGIGT 358
QY 349 SSVNKTIVVEYQRENERTANLOSMTNIDTKNHTVEQTIYINPLRSKAKETNV----- 402
Db 359 NTSKTYLIDYEKGGQHNLSIKGTIDQIDKTNNTYQTIYVNF-----SGDNVVLPAIT 413
QY 403 -NISNGDESGTIIIDSTIIKVKYVGNQNLPSNRITYDYSEYEDVTNDYAOIGNNNDV 461
Db 414 GNLIPTKSNALIDAKNTDIAKYRV-DNANDLSESYVNSDFEDVTNQVRISFPNANQY 472
QY 462 NINF---GNIDSPYIIKVIISKYDPNKKDYTTTIOQVTMTQTIINEYTGEE--RTASYDNT 515
Db 473 KVEFPDDDDQITTPYIVVNGHIDPASTG-----DLALSTFYGYDSNFIKRSMSWDNE 526
QY 516 IASTSSGOGG-DLP--PEKTYKIGDY--VWEDVDKD-GIONTNDN 556
Db 527 VAFNNGSGDGDIDKPVVPEQPEGEIEPIPEDSDSPSGSDS 573

RESULT 15
Q99VJ4 PRELIMINARY; PRT; 989 AA.
AC Q99VJ4;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE FIBRINOGEN-BINDING PROTEIN A, CLUMPING FACTOR.
GN CLFA OR SA0742.
OS Staphylococcus aureus (strain N315).
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OX Bacillus/Staphylococcus group; Staphylococcus.
RN NBI_TaxID=158879;
RP SEQUENCE FROM N.A.
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.I., Nagai Y., Iian J., Ito T., Kanamori M.,
RA Matsumaru H., Maruyama A., Murakami H., Hosoyama A., Mizutani-Ui Y.,
RA Takahashi N.K., Sawano T., Inoue R.I., Kaito C., Sekimizu K.,
RA Hiraoka H., Kuhara S., Goto S., Yabuzaki J., Kanehisa M.,
RA Yamashita A., Oshima K., Furuya K., Yoshino C., Shibata T., Hattori M.,
RA Ogasawara N., Hayashi H., Hiratsugu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RL aureus".
RL Lancet 357:1225-1240(2001).
EMBL: AP003131; BAB41975.1; -.
InterPro: IPR000515; BPD_transp.
InterPro: IPR001899; Gram_pos_anchor.
PROSITE: PS00402; BPD_TRANSP_INN_MEMBER, UNKNOWN_1.
KW PROSITE; PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.
SQ complete proteome.
SQ SEQUENCE 989 AA; 102407 MW; DA6E807539623467 CRC64;

```

Query Match 14.6%; Score 452; DB 16; Length 989;
 Best Local Similarity 25.9%; Pred. No. 7.7e-11;
 Matches 152; Conservative 109; Mismatches 224; Indels 102; Gaps 27;

```

QY 8 SDEKNDVYNNQSIINTDNNQIIEKEETNNYDGIKRESDRETTSTTNDENEATPLOG 67
Db 51 SNEKSNDSSVSAAPRTDFTN-----SDTKTSNTN--NGETISVAG 91
QY 68 TPQDNTHLIEEYKSSSVSSNSIDTAQPSHTTINREESVQTSNVEDSHVDFANS 127
Db 92 NP-----AQOETQSSSTNATTEETPVTGEATTTTNOANTPATQS-SNTNAEELVN- 143
QY 128 KIKESNTSGKEKNTIQPKVKEDSTTSQPSGTINDEKISNODELNL-LPINEYENKA 166
Db 144 --QTSNETHNDNTNIV-----SSVNSPQNSTNAENVTSTQDTSTTEATPSN--NES 189

```

```

QY 187 RPLSTTSA-----QPSIKRVTVNQLAA-----EAGSNVHLIKVTDQSTTEGYD 231
Db 190 APQNTDASNDQVYSAVNPSTPRKRAFSLAAVADAPAGTDITN--QLTDVAKY--ID 244
QY 232 SEGVIKADNENLIYDVEYDVKVSGDTMTYDIDKNVPSDLTDSFTIPKIDNSGEI 291
Db 245 SGTIVPHQAGYKLVNGFSPNSAVKGDFTKIVPEKELNLTGVTSTAKVPPIMAGD-QV 303
QY 292 IATGYDNKKKOITFTDQVVDKXENIKAKLKLTSYIDSKVPNNKTL-DVEYKTL-- 348
Db 304 LANGVIDS-DGNVIYFTDYDNKENVTANTMPAYID----PENVTKGAVTLTGIGT 358
QY 349 SSVNKTIVVEYQRENERTANLOSMTNIDTKNHTVEQTIYINPLRSKAKETNV----- 402
Db 359 NTSKTYLIDYEKGGQHNLSIKGTIDQIDKTNNTYQTIYVNF-----SGDNVVLPAIT 413
QY 403 -NISNGDESGTIIIDSTIIKVKYVGNQNLPSNRITYDYSEYEDVTNDYAOIGNNNDV 461
Db 414 GNLIPTKSNALIDAKNTDIAKYRV-DNANDLSESYVNSDFEDVTNQVRISFPNANQY 472
QY 462 NINF---GNIDSPYIIKVIISKYDPNKKDYTTTIOQVTMTQTIINEYTGEE--RTASYDNT 515
Db 473 KVEFPDDDDQITTPYIVVNGHIDPASTG-----DLALSTFYGYDSNFIKRSMSWDNE 526
QY 516 IASTSSGOGG-DLP--PEKTYKIGDY--VWEDVDKD-GIONTNDN 556
Db 527 VAFNNGSGDGDIDKPVVPEQPEGEIEPIPEDSDSPSGSDS 573

```

Search completed: July 30, 2002, 10:40:30
 Job time: 569 sec

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)